Author:
Watson, Jessica C

Title:
The diagnostic utility of inflammatory markers in primary care

a mixed methods study

The diagnostic utility of inflammatory markers in primary care: a mixed methods study

JESSICA WATSON

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy (PhD) in the Faculty of Health Sciences. Bristol Medical School. March 2021.

Word Count: 71,730

ABSTRACT

Inflammatory markers are non-specific blood tests including C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) and plasma viscosity (PV). These tests are commonly used by general practitioners (GPs) as an aid to the diagnosis and monitoring of inflammatory conditions including infections, autoimmune conditions, and cancers. Previous research into inflammatory markers has mostly been based in secondary care and explored relationships with single disease outcomes. This means it is less useful for GPs, who often use inflammatory markers in patients with undifferentiated symptoms who have multiple possible disease outcomes. There is also a lack of information about how GPs communicate the rationale for testing and the meaning of results with patients. My research addresses these evidence gaps.

Firstly, I conducted a series of quantitative studies using electronic health records of nearly 200,000 patients from the Clinical Practice Research Datalink (CPRD). The aim was determine the epidemiology and diagnostic utility of inflammatory marker tests in primary care for relevant disease: defined as any infection, autoimmune disease or cancer. Secondly, I conducted a qualitative study, completing a total of 80 interviews with patients before and after receiving inflammatory marker results, and with the GPs who requested these tests. The aim was to explore the meaning of inflammatory markers for doctors and patients.

I have shown that, contrary to GPs perceptions, inflammatory markers are not a useful 'rule-out' test: in fact they miss around half of relevant disease in primary care. Testing more than one inflammatory marker simultaneously does not increase diagnostic accuracy. Although these tests are often used for reassurance, patients with ongoing symptoms perceived that normal results were unhelpful. I identified a lack of shared understanding and multiple barriers to shared decision-making. Results will be of interest to patients, GPs and commissioners wanting to optimise inflammatory marker testing in primary care.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED:……*Jessica Watson*……..DATE……26/03/2021…….

Acknowledgements

I would like to start by thanking my outstanding supervisory team. I am so fortunate to have met Willie Hamilton back in 2005 when I was a newly qualified doctor dabbling in research. He has championed and supported me ever since and inspired my interest in academic primary care. I feel so lucky to have been welcomed by his whole team on my regular trips to Exeter and have gained so much by my affiliation with the Exeter CanTest group. Chris Salisbury has provided oversight to the whole thesis, kept me on track, stayed patient with my constant dabbling and distractions and always given me honest and wise advice. Penny Whiting has provided razor sharp methodological expertise, has helped develop me as a diagnostic researcher and helped me to think beyond inflammatory markers and grow into a wider field of diagnostics. Jon Banks has guided me through unfamiliar territories of qualitative research and helped me build and develop the PPI component of my thesis.

I would like to thank my patient and public participation group for their invaluable advice and steering at key points in the thesis. I would also like to express my gratitude to all the GPs and patients who allowed me to interview them for the qualitative study and shared their experiences and wisdom. I am also thankful to all the support from practice managers, phlebotomists and healthcare assistants without whom it would have been impossible to recruit to my study.

I would like to thank the many colleagues who have helped me along the way: Polly Duncan, Lucy Pocock, Lucy Potter, Sam Merriel for reviewing endless boring code-lists, Sam Creavin for advice on navigating the challenges of a PhD as a clinical academic, my office roomies Natalie Blencowe, Barry Main and Rebecca Barnes for sharing coffee machines and words of wisdom (at least until Covid disruption), and my writing day buddies especially Anne Herbert for virtual support during my final write up.

I am of course indebted to the NIHR for funding this work through a Doctoral Fellowship (DRF-2016-09-034) and the NIHR ARC West for adopting my PhD and supporting PPI and dissemination. I am hugely grateful to Dr Peter Brindle at BNSSG CCG for his enthusiasm in my work and for linking me with Becca Robinson, Clinical Effectiveness Programme Manager at the CCG who helped me to move out of the 'ivory tower' into the world of knowledge mobilisation. I would like to thank Debbie Sharp who since my very first day in Academic Primary Care back in 2005 has supported and inspired me.

I am so grateful to my Mum, who let me use her garden room when working from home became too much, and for providing endless cups of tea and coffee, cakes, and delicious food. I would like to thank my Dad who has always been my biggest fan and helped me to believe in myself. I would like to thank both my parents for proof-reading.

Above all I would like to thank my husband Peter Watson for being the most amazing husband and Daddy to our four wonderful children: Saskia, Guy, Fred, and Pearl. As if that wasn't enough, he was also mad enough to suggest that our dog, Lottie, have puppies around the time of my thesis write up. Our eight golden retriever puppies deserve special thanks for simultaneously driving me mad and keeping me sane during the dark winter days of covid lockdown.

List of publications

The research presented in this thesis has been published in the following peer-reviewed papers:

- Watson J, Nicholson BD, Hamilton W, & Price S. (2017) Identifying clinical features in primary care electronic health record studies: methods for codelist development. *BMJ Open* 7(11) https://doi.org/10.1136/bmjopen-2017-019637
- Watson J, Salisbury C, Whiting P, Banks J, Pyne Y, & Hamilton W. (2019). Added value and cascade effects of inflammatory marker tests in UK primary care: a cohort study from the Clinical Practice Research Datalink. British Journal of General Practice, 69(684), e470-e478. https://doi.org/10.3399/bjgp19X704321
- Watson J, Jones H, Banks J, Whiting P, Salisbury C, & Hamilton W. (2019). Use of multiple inflammatory marker tests in primary care: using Clinical Practice Research Datalink to evaluate accuracy. British Journal of General Practice, 69(684), e462-e469. https://doi.org/10.3399/bjgp19X704309
- Watson J, Salisbury C, Banks J, Whiting P, & Hamilton W. (2019). Predictive value of inflammatory markers for cancer diagnosis in primary care: a prospective cohort study using electronic health records. British Journal of Cancer, 120(11), 1045-1051. https://doi.org/10.1038/s41416-019-0458-x
- Watson J, Whiting P, Salisbury C, Banks J, & Hamilton W. (2020). Raised inflammatory markers as a predictor of one-year mortality: A cohort study using primary care electronic health record data. BMJ Open. https://doi.org/10.1136/bmjopen-2019-036027

In addition to these core papers which are direct output from this thesis, I have also published the following affiliated peer-reviewed papers during my fellowship, which would not have been possible without the methodological training in diagnostic test accuracy research gained through my doctoral research:

- Mytton O, McMarthy N, Watson J, Whiting P. (2021). Interpreting a lateral flow SARS-CoV-2 antigen test. BMJ (submitted)
- Watson J, Whiting PF, & Brush JE. (2020). Interpreting a covid-19 test result. BMJ, 369:m1808. https://doi.org/10.1136/bmj.m1808
- Watson J, Richter A, & Deeks J. (2020). Testing for SARS-CoV-2 antibodies. BMJ, 370:m3325. https://doi.org/10.1136/bmj.m3325
- Martin J, Watson J, & Barnes R. (2020). Shared decision making about blood tests: secondary analysis of video-recorded primary care consultations. British Journal of General Practice, 70(694):e339-e347 https://doi.org/10.3399/bjgp20X709409
- Watson J, Mounce L, Bailey S, Cooper S, & Hamilton W. (2019). Rational Testing: Blood markers for cancer. BMJ, 367, l5774. https://doi.org/10.1136/bmj.l5774
- Watson J, de Salis I, Banks J, Salisbury C. (2017) What do tests do for doctors? A qualitative study of blood testing in UK primary care. Family Practice. 34(6):735-739. https://doi.org/10.1093/fampra/cmx051

I have also published the following letters which are affiliated to my Doctoral Research Fellowship:

- Watson J, Salisbury C, Banks J, Whiting P, Hamilton W. (2019) High results (letter). BJGP. https://bjgp.org/content/69/684/e470/tab-e-letters#high-results

- Watson J, Hamilton F, Bailey S, Mounce L, Hamilton W. (2019) Clinical implications of increased testing in primary care (letter). BMJ. https://www.bmj.com/content/364/bmj.l175
- Watson J. (2017) All tests can sometimes cause more harm than good (letter). BMJ. 358:j4070 https://doi.org/10.1136/bmj.j4070

Table of Contents

List of figures

List of tables

List of Abbreviations

AUC	Area under the receiver operating curve
BMJ	British Medical Journal
BJGP	British Journal of General Practice
CI	Confidence interval
CRP	C-reactive protein
CPRD	Clinical Practice Research Datalink
DOR	Diagnostic odds ratio
ESR	Erythrocyte sedimentation rate
GP	General Practitioner
HES	Hospital episode statistics
IMD	Index of multiple deprivation
NICE	National Institute for Health and Care Excellence
NCRS	National Cancer Registry Service
NPV	Negative predictive value
OR	Odds Ratio
ONS	Office for National Statistics
PIL	Participant information leaflet
PPI	Patient and public involvement
PPV	Positive predictive value
PV	Plasma viscosity
POC	Point of care
RCT	Randomised controlled trial
ROC	Receiver operating curve
SAUC	Summary area under receiver operating curve

CHAPTER 1. BACKGROUND

1.1. Introduction

This thesis explores the diagnostic utility and clinical practice of inflammatory marker testing in primary care, using a mixed methods approach. Inflammatory markers are non-specific tests which are commonly used in primary care as an aid to the diagnosis and monitoring of inflammatory conditions including infections, autoimmune conditions, and cancers.

My interest in this research area has arisen from my own clinical experiences as a General Practitioner (GP), and the challenges I have experienced when using inflammatory marker tests, including uncertainty about when to test, how to interpret results, and how to share these with patients. In 2012 I was first author of a British Medical Journal (BMJ) review article on inflammatory marker testing; whilst I found many studies of inflammatory marker tests from secondary care, I was surprised by the lack of research in primary care settings.(1) Most of the published research focussed on single disease outcomes, which did not reflect the reality of my experience in primary care, where patients with undifferentiated symptoms had multiple possible disease outcomes. To explore the way inflammatory markers were really being used in primary care, I conducted a qualitative research project, interviewing primary care clinicians about their experiences of inflammatory marker testing in primary care, during my pre-doctoral Academic Clinical Fellowship.(2) The issues arising from these qualitative interviews were used to help generate the research questions for this thesis. In particular, GPs described using inflammatory markers in patients with non-specific symptoms as a generic test to rule-out serious underlying disease and to provide reassurance to patients and GPs. This practice was largely unsupported by evidence and guidelines. GPs perceived that normal inflammatory markers were useful for reassurance, but there was a lack of

evidence on patients' perspectives of testing. GPs described their decision-making for inflammatory marker testing as encompassing multiple non-medical reasons for testing, and their accounts of diagnostic decision-making and information sharing did not reflect the ideals of shared decision-making.

I therefore set out to look at inflammatory marker testing using a different approach to previous researchers. Firstly, I wanted to explore the diagnostic accuracy of inflammatory markers for overall relevant disease including infections, autoimmune conditions, and cancers, to determine their utility as a non-specific 'rule-out'. Secondly, I wanted to use qualitative methods to explore the communication and shared understanding of inflammatory marker tests between doctors and patients.

1.2. What are inflammatory markers?

Inflammatory markers are blood tests which are commonly used in clinical practice to detect inflammation in the body, which may occur due to infections, autoimmune conditions, or some cancers. The three main inflammatory markers used in primary care are C reactive protein (CRP), erythrocyte sedimentation rate (ESR) and plasma viscosity (PV). Procalcitonin is a newer inflammatory marker, which is now used routinely in hospital and A+E settings, but has yet to be introduced to routine primary care in the UK. For the purposes of this thesis, I focus on the three main inflammatory markers: CRP, ESR and PV.

The ESR is the oldest of these three inflammatory markers and is measured as the distance in millimetres that erythrocytes settle in anticoagulated whole blood in one hour. Therefore, it is not a measure of a single acute phase protein, and is dependent on a number of other factors, including anaemia or polycythaemia, delays between sample taking and lab processing, medications (e.g. oral contraceptive pill and heparin), sex, age, pregnancy, ethnicity and obesity.(3) As

a result of these limitations, many laboratories and guidelines are moving away from the use of ESR, except in a limited number of specific conditions.(4)

Plasma viscosity is measured by calculating the force needed to send plasma down a thin tube in a given time, which is associated with the amount of protein in the blood. Thus, it is raised with elevations in acute phase proteins and with increased paraproteins in the blood produced by certain tumours. It can be considered technically superior to ESR, being unaffected by anaemia or polycythaemia, or by delays between sampling and measurement, and has results which are independent of age or sex.(5) However it can be technically challenging to perform and is only offered by some laboratories in the UK. Many of the factors that determine ESR and plasma viscosity have long half-lives, so elevated results may be due to an event which happened weeks to months previously and may have resolved at the time of measurement. This makes them less suitable for assessing acute changes and monitoring response to treatment and means that results may be difficult to interpret.

CRP is a pentameric protein molecule first described in 1930, which is produced by hepatocytes, and named due to its reaction with pneumococcal C-polysaccharide.(6) CRP binds with phosphocholine, then interacts with antibody receptors on phagocytic cells to facilitate phagocytosis. CRP therefore has a role in clearing pathogens in infection, plus damaged, necrotic or apoptotic cells from other types of inflammatory response. CRP has a more rapid response and shorter half-life than ESR and plasma viscosity, allowing it to be used to detect and monitor the progress of systemic inflammation in the body.(7) In ostensibly healthy subjects, CRP tends to increase slightly with age,(8) and is also increased with obesity.(9)

Most inflammatory marker tests are performed on venous blood samples in laboratory settings; however, CRP tests are also available as point of care tests, using finger-prick blood samples. Point of care CRP tests are increasingly available in the UK, and offer the potential to provide rapid feedback, allowing

test results to inform immediate clinical decision-making within a general practice consultation.

1.3. Epidemiology of inflammatory marker testing

1.3.1. Temporal trends in inflammatory marker testing rates

Rates of inflammatory marker testing in primary care are rising in the UK. A recent paper by O'Sullivan *et al* analysed all tests conducted between April 2000 and March 2016 in the Clinical Practice Research Datalink (CPRD), a large database covering about 7% of the UK population.(10) They found primary care CRP testing rates showed a consistent linear increase from 2000-2015 from 91.6 tests per 10,000 age and sex adjusted person years in 2000 to 924.4 tests per 10,000 person-years in 2015, an average increase of 16.5% per annum. Only five blood tests - vitamin D, ferritin, troponin, folate and B12 - had higher annual increases than CRP. Use of ESR also increased, though less steeply, from 501.1 tests per 10,000 person years in 2000 to 1027.2 per 10,000 in 2015, an average annual increase of 4.7%. Use of plasma viscosity was not examined in this study, although another smaller study using CPRD did examine plasma viscosity, which increased by 31% from 129 to 170 tests per 10,000 person years between 2005 and 2009.(11)

These large increases in testing rates seen over time have not been matched by concomitant rises in rates of diagnosis of disease. Presumably therefore testing is shifting into populations which are at a lower risk of disease.

1.3.2. **Variation in inflammatory marker testing rates**

Variation in rates of inflammatory marker testing between GP practices was examined in a linked paper by O'Sullivan *et al*, using a subset of the same dataset.(12) They calculated a co-efficient of variation for 44 primary care tests to identify variation in testing rates between GP practices. CRP was one of only seven tests which showed both high variation in testing rates and high use. The observed differences could potentially reflect different rates of pathology in different practices, which could warrant different levels of testing. However, the differences persisted after adjustment for practice level age distribution, sex and deprivation. Residual confounding due to unmeasured differences between practices could be an issue, however it seems unlikely that this accounts for all the remaining variation seen. Some of this variation is therefore likely to reflect differences in clinicians' habits and thresholds for testing. This is in keeping with my pre-doctoral qualitative research, in which clinicians reported they felt uncertain when to test inflammatory markers, with decisions about testing based on experience and habit rather than guidelines or evidence:(2)

> *With GPs, a lot of things are oh well it is what I do and it's what I have always done and erm it seems to work, so yeah I'll just do it a bit more … so what we end up doing is kind of making it up.*

Another paper described the variation in the proportion of abnormal CRP and ESR test results, across practices in Oxfordshire in 2016.(13) This showed that the annual unadjusted proportion of abnormal CRP tests per practice varied from 22% to 42%. For ESR the variation in proportion of abnormal results was greater still, varying from 12% to 41%. After adjustment for practice list size, age, sex and gender and index of multiple deprivation, twelve (17%) and seven (10%) practices were greater than 3 standard deviations below the mean proportion of abnormal test results for CRP and ESR respectively. This suggests that not only do rates of testing vary, but thresholds for testing vary between practices, with some practices testing behaviour yielding fewer abnormal results. This is in keeping with the hypothesis that high testing rates in some practices cannot

solely be explained by differences in rates of pathology. Similar variation in rates of inflammatory marker testing rates have been found in studies in Spain(14, 15) and Sweden.(16)

1.3.3. Potential harms of inappropriate testing

A major limitation of the research described above is that none of it directly explores the 'appropriateness' of testing, by measuring reasons for testing and outcomes of testing. There is a long history of using geographical variation and between practice variation to highlight potential overuse and underuse of healthcare resources,(17) for example the NHS Atlas of Variation.(18, 19) Within this wider literature, factors predicting variation in testing rates have been extensively studied, both at GP practice level,(20) and also by looking at characteristics of clinicians which are associated with higher or lower rates of testing.(21) This is often interpreted as 'unwarranted' variation, however this evidence must be interpreted with caution, particularly if used for feedback to individual practices or clinicians, as there may be good reasons for variation in testing rates which are not measured or measurable. Nonetheless, the triangulation of evidence showing rising testing rates, variation in rates of testing and variation in rates of abnormal results, combined with qualitative evidence demonstrating uncertainty surrounding inflammatory marker testing,(2) all points to potential inappropriate use of inflammatory marker tests.

If tests are used inappropriately, either through over-testing or undertesting, this influences the interpretation, and usefulness, of the results. This is because tests work in a Bayesian fashion. This means that the interpretation of any test depends on two things: the performance characteristics of the test itself and the subgroup of people it is performed on (the prior odds or pre-test probability of disease).

No test is perfect - test accuracy is measured by two parameters: sensitivity or true positive rate, and specificity or true negative rate. These measures are

relatively stable. However, the clinically more important parameters - positive predictive value and negative predictive value, vary depending on pre-test probability.

This can have counter-intuitive implications: for example – studies have shown that people with normal test results have an increased risk of cancer.(22) This is because the mere fact that a test has been conducted predicts cancer and this additional risk is only partly eliminated by a negative test result. The higher the pre-test probability of disease, the more likely it is that these false negatives will occur. Missed diagnosis of cancer is a particular concern, with much of the UK's poor record in cancer outcomes blamed on diagnostic delays. If raised inflammatory markers are not investigated appropriately this may represent a missed opportunity for early diagnosis of cancer or other serious disease. However, conversely, if inflammatory markers have low sensitivity for cancer, and are used on those at high risk of disease, it is equally possible that false negatives could lead to false reassurance.

Given the rising rates of testing, the opposite problem over-testing and false positives is more likely to be a concern.(23, 24) As testing rates increase, in lower risk populations, this leads to a reduction in the signal-to-noise ratio, leading to lower positive predictive values and increased false positives. These false positives may cause anxiety for patients, and may lead to further tests, appointments and referrals, in what Deyo has called the 'cascade effect of medical technology'.(25) Recent estimates have suggested that more than half of abnormal results from laboratory tests ordered by family physicians could be false positives.(26) Although this concept of cascade testing has been around for decades,(25) it is rarely measured,(27) and the overall frequency and implications of cascade testing on primary care workload is unknown. My qualitative research showed that GPs perceive cascade testing to be a particular issue with inflammatory markers, because of the non-specific nature of these tests.(2) As well as patient harms, this could lead to increased workload for GPs(28) who are already overstretched.(29) The carbon footprint of pathology testing is also

significant,(30) and an increasing priority, given current targets for reducing waste in healthcare(31) and the NHS target to achieve net zero emissions by 2040.(32)

Although the unit cost of inflammatory marker tests is relatively low, overall costs of testing are considerable; in 2014 at North Bristol NHS Trust over 120,000 primary care requests for inflammatory markers were performed at a cost of £177,000, for a population of 500,000. These costs do not include the substantial additional costs of phlebotomists', GPs' and patients' time and the potential costs generated by further cascades of tests and follow up appointments.

1.4. Quantitative evidence of inflammatory marker test accuracy

1.4.1. Reasons for inflammatory marker testing

Inflammatory markers are non-specific tests which detect inflammation in the blood but do not indicate the cause of this inflammation. Diseases with prominent activation of the inflammatory response fall into three main groups: infections, autoimmune conditions and some malignancies.

Inflammatory markers can be measured by primary care clinicians for four main possible reasons:

- Diagnosis of specific disease (see 1.4.2)
- Monitoring of disease progression or treatment response (see 1.4.3)
- Non-specific testing for systemic disease (see 1.4.4)
- Testing to predict future outcomes or prognosis (see 1.4.5)

Several studies, now decades old, have looked at the reasons for general practitioners using inflammatory markers. They found that inflammatory

markers were used for a range of indications, with about 44-47% requested for specific diagnostic purposes, 27-33% for monitoring disease, and 14-28% for non-specific diagnostic purposes. (33, 34) Testing to predict future outcomes is mostly of research interest rather than having direct clinical implications but may have future implications for inflammatory marker testing rates and is therefore discussed briefly in section 1.4.5.

1.4.2. Testing for diagnosis of specific diseases

I was first author of a literature review published in the BMJ in 2012 which summarised the evidence for using inflammatory marker tests in diagnosis.(1) Since this was published there has been a significant amount of further research, so I have updated the summary table from my BMJ paper to include more recent systematic reviews of the accuracy of inflammatory markers for diagnosis (**Table 1**). Most of this evidence focuses on single disease outcomes, mostly infections, and the majority is based in secondary care; positive predictive values will be lower in primary care where disease prevalence is lower. I will consider this evidence in more depth in relation to the three main categories of disease below: infections, autoimmune conditions and cancers. These three categories of disease have been chosen as they were the main disease groups identified by my literature review of inflammatory marker testing.(1) The combination of infection, autoimmune conditions or cancers will be referred to throughout this thesis as 'relevant disease'.

Infections

Most inflammatory marker research into infections uses CRP as the index test for detecting and monitoring acute infections, due to its shorter half-life compared to PV and ESR. In secondary care settings CRP has been studied as a diagnostic aid for multiple types of infection including endocarditis, meningitis, peritonitis,

sepsis, bone and joint infections, chorioamnionitis, post-operative infections, and more recently SARS-CoV-2 infections (**Table 1**). Many of these infective outcomes are relatively rare in primary care settings.

Research in primary care settings has mostly focussed on the utility of CRP tests for the diagnosis of acute respiratory infection. Systematic reviews of diagnostic accuracy studies in primary care settings have been published for community acquired pneumonia (35) and acute sinusitis.(36) Inflammatory markers have also been studied in primary care settings to assess their diagnostic utility in predicting serious infection in children with a fever,(37) and to predict infective exacerbations of chronic obstructive pulmonary disease.(38)

In these clinical scenarios, CRP has been found to have only modest diagnostic accuracy, which is generally insufficient to safely rule acute infection either in or out. As a result there has been significant interest in developing clinical prediction rules incorporating CRP test results with relevant clinical features, including prediction rules for serious infection in children with fever,(39) and community acquired pneumonia.(40) Clinical prediction rules for the diagnosis of community acquired pneumonia have been extensively investigated, with meta-analysis of individual patient data demonstrating improved discrimination using CRP tests in addition to symptoms and signs as measured by the area under curve (AUC).(41) The delay in obtaining the results of inflammatory marker test results if the specimen requires analysis off site contributes to slow uptake of these clinical decision rules in mainstream UK clinical practice. With point of care (POC) testing widespread in Europe, and entering clinical practice in primary care in the UK, this may lead to changes in the use of inflammatory markers in the future.

The research into the diagnostic accuracy of CRP for community acquired pneumonia mostly uses radiological changes on chest x-ray as the reference standard; however clinical prognosis of acute respiratory infection is arguably more important, as antibiotics should ideally be targeted to those with the highest risk of complications, which may not correspond to radiological

features.(42) To address this, Bruyndonckx *et al* developed a clinical prediction rule for adverse outcome (defined as re-attendance in primary care or hospital admission) using symptoms, signs and tests in six European countries; the addition of CRP to the model based on symptoms and signs did not improve prediction.(43)

Despite this, there has been significant interest in the utility of POC CRP tests to reduce primary care antibiotic prescribing rates, driven by concerns about rising rates of antimicrobial resistance. Multiple randomised controlled trials (RCTs) have been performed, with a 2014 Cochrane review concluding that use of POC CRP tests leads to overall reduction in antibiotic prescribing for suspected lower respiratory tract infections.(44) Similar reductions in antibiotic prescribing have also been demonstrated for infective exacerbations of chronic obstructive pulmonary disease.(45) The apparent contradiction between RCTs findings and diagnostic test accuracy studies may reflect the fact that point of care CRP tests exert some of their effect as a behavioural intervention to improve patients' and doctors' confidence in prescribing decisions. Longer term follow up of some RCTs has shown that this reduction in antibiotic prescribing with POC CRP testing wanes with time.(46) Qualitative studies have demonstrated a number of barriers to implementation(47) which may explain the lack of use of POC CRP testing in UK primary care, despite UK National Institute for Health and Care Excellence (NICE) guidelines suggesting GPs 'consider' their use in suspected community acquired pneumonia.(48)

This thesis focuses on the overall utility of inflammatory markers for overall relevant disease, rather than strengthening this evidence base for specific infective indications. However, if (or when) point-of-care CRP tests become widely available in the UK they could end up being used for indications beyond those for which they have been evaluated. This has been demonstrated in Norway where GPs have been found to use point-of-care CRP tests for skin, digestive, ear, eye and 'general unspecified' symptoms.(49) Further research to

determine an optimal evidence-based approach to the interpretation and management of raised inflammatory markers is therefore timely.

Autoimmune conditions

Inflammatory markers are established first-line tests for a small number of autoimmune diseases. The classic conditions are polymyalgia rheumatica or giant cell arteritis. There are no specific serological tests for these conditions, so diagnosis is generally made on the basis of symptoms and signs of myalgia, headache and systemic features, alongside a raised inflammatory marker.(50, 51) Evidence mostly comes from secondary care settings, and generally requires a temporal artery biopsy as the reference standard, so patients evaluated would be likely to have more severe disease than primary care populations.(50, 51) Studies show that both ESR and CRP can sometimes be normal in polymyalgia and giant cell arteritis;(52, 53) UK guidelines therefore recommend both tests for diagnosis.(54)

Diagnosis of inflammatory arthritis is another common clinical reason for inflammatory marker testing. NICE guidelines state that urgent referral of patients with clinical evidence of rheumatoid arthritis should not be delayed even if inflammatory marker tests are normal:(55) this is based on evidence that 35% to 45% of patients with rheumatoid arthritis have normal inflammatory markers at diagnosis.(56)

Inflammatory markers have been studied as an aid to diagnosis of inflammatory bowel disease; however, the newer test, faecal calprotectin, has largely superseded inflammatory marker blood testing in the diagnosis of inflammatory bowel disease and is now recommended as a first line test by NICE.(57) The NICE diagnostics guidance for faecal calprotectin concludes that CRP and ESR have poor sensitivity for inflammatory bowel disease and states '*there therefore seems little point in doing [CRP and ESR] even if calprotectin was not available*'.(57)

Cancer

Cancer diagnosis in primary care can be challenging; many of the early symptoms are non-specific and can be difficult to differentiate from the symptoms of common benign conditions. GPs need to triage patients with low-risk symptoms to identify those needing further investigations, using additional 'clues' from history, examination and investigations.(58) One triaging tool increasingly used in clinical practice is inflammatory marker testing. These are often performed as a 'rule-out' test by clinicians trying to exclude serious underlying disease, including cancer.(2) This practice is largely unsupported by evidence and inflammatory markers are not recognised within current guidelines for cancer diagnosis. The main exception to this is multiple myeloma, with a recent large case-control study in primary care confirming the diagnostic utility of ESR and PV; CRP had lower diagnostic accuracy and was therefore not recommended by the authors.(59) This is reflected in the NICE cancer guidelines, which recommend ESR or PV as first line tests for suspected myeloma. Evidence for inflammatory marker testing in primary care for other types of cancer is limited, with case-control studies for pancreatic cancer,(60) bladder cancer,(61) renal cancer(62) and non-Hodgkin's lymphoma,(63) showing associations, but with positive predictive values <1%; too small to be clinically useful. Cohort studies in the general population (irrespective of symptoms) have shown an association between CRP and risk of future cancer,(64) although not strong enough to be clinically useful for identification of symptomatic cancer. There have been no previous primary care studies to measure the predictive value of inflammatory markers for overall cancer diagnosis. Current NICE guidelines recommend urgent referral or investigations for anyone with a cancer risk of ≥3%,(65) so even relatively low predictive values could potentially be clinically useful to improve cancer diagnosis.

Table 1: *Systematic reviews of the accuracy of inflammatory markers for diagnosis of specific conditions*

Target condition (test)	Setting	Study type	Outcome*
Acute rhinosinusitis (CRP, ENT)	Primary care and ENT clinics	Systematic review (4 reports of CRP/ESR; 789 patients)	CRP summary sensitivity 0.34 (0.21 to 0.51), specificity 0.88 (0.79 to 0.94). ESR summary sensitivity 0.43 (0.29 to 0.58), specificity 0.83 (0.70 to 0.92) (36)
Bacterial chest infections in adults (CRP)	Primary care; accident and emergency departments	Systematic review (8 reports; 1230 patients)	Likelihood ratio of raised CRP 2.1 (1.8 to 2.4); negative likelihood ratio 0.33 (0.25 to 0.43).(35) Similar results in an earlier review (66) and a subsequent study.(67)
Bacterial chest infection in children (CRP, ESR)	Secondary care	Systematic review (8 reports; 1230 patients)	Pooled odds ratio for raised CRP and bacterial infection 2.6 (1.2 to 5.6) (68)
Infective exacerbations in patients with COPD	Secondary care	Systematic review (59 reports investigating 134 biomarkers; 28 reports of CRP)	Due to heterogeneity of studies no meta-analysis performed.(69)
Infective exacerbation in patients with cystic fibrosis (CRP, ESR)	Secondary care	Systematic review (18 reports of 35 biomarkers; 9 reports of CRP; 3 reports of ESR)	Due to heterogeneity, no meta-analysis performed. (70)
Active pulmonary tuberculosis (CRP)	Secondary care	Systematic review (9 reports; 1793 patients)	Sensitivity 0.93 (0.88 to 0.98) specificity 0.60 (0.40-0.75) among out-patients (71)
Appendicitis in children with abdominal pain (CRP)	Secondary care (mainly emergency departments)	Systematic review (5 studies of CRP, 1 of ESR; 730 and 162 children respectively).	Likelihood ratio for CRP>35mg/L 5.2 (1.7 to 16). For ESR>20mm/h, 3.8 (1.8 to 8.1)(72)
Appendicitis in adults with abdominal pain (CRP)	Secondary care	Systematic review (7 studies; 1011 patients)	Summary sensitivity 0.57 (0.39 to 0.73), specificity 0.87 (0.58 to 0.97).(73)
Serious conditions in abdominal pain (CRP)	Secondary care (emergency departments)	Systematic review (3 studies, 2961 patients)	Sensitivity 0.77 (0.74 to 0.79), specificity 0.61 (0.59 to 0.64) at cut-off CRP>10mg/L (74)
Intra-abdominal infection after elective colorectal surgery (CRP)	Secondary care	Systematic review (11 studies, 2692 patients)	Pooled sensitivity 0.75 (0.69 to 0.81), specificity 0.72 (0.70 to 0.74).(75) Similar findings in one other study (76)
Bacterial peritonitis in patients with cirrhosis (CRP)	Secondary care	Systematic review comparing procalcitonin with CRP (7 reports of CRP; 499 patients)	Pooled sensitivity 0.76 (0.58 to 0.88), specificity 0.81 (0.63 to 0.92) (77) Similar results in an earlier review (78)
Complications post abdominal surgery (CRP)	Secondary care	Systematic review (7 studies; 1986 patients)	Summary sensitivity 68%, specificity 72%. (79) Similar findings in two other reviews (80) (81)
Sepsis in adults (CRP)	Secondary care	Systematic review (9 reports; 495 sepsis patients, 873 non-sepsis patients)	Pooled sensitivity 0.80 (0.63 to 0.90) specificity 0.61 (0.50 to 0.72). (82)
Neonatal sepsis (CRP)	Secondary care	Systematic review (28 reports; 2661 patients)	Pooled sensitivity 0.71 (0.63 to 0.78) specificity 0.88 (0.80 to 0.93)(83)

Late onset infection in newborn infants (CRP)	Secondary care	Cochrane systematic review (20 reports; 1615 infants)	At median specificity (0.74), sensitivity was 0.62 (0.50 to 0.73) (84)
Bacterial meningitis in children (CRP)	Secondary care	Systematic review (8 studies, 616 patients)	Pooled sensitivity 0.70 (0.64 to 0.76), pooled specificity 0.83 (0.79 to 0.87) (85)
Bacterial meningitis in adults (CRP)	Secondary care	Systematic review of procalcitonin versus CRP (7 studies of CRP; 635 patients)	Pooled sensitivity 0.92 (0.75 to 0.88) specificity 0.81 (0.77 to 0.84). (86)
Bacterial infection in adults with febrile neutropenia (CRP)	Secondary care	Systematic review (13 studies; 1712 patients)	Pooled sensitivity 0.77 (0.62 to 0.87), specificity 0.58 (0.44 to 0.71) (87)
Adverse outcomes in febrile neutropenic episodes in children (CRP)	Secondary care	Systematic review (7 studies; 731 episodes)	Pooled sensitivity 0.65 (0.41 to 0.84), specificity 0.73 (0.63 to 0.82). (88) Updated study showed procalcitonin may have better discriminatory ability. (89)
Pyelonephritis in children (CRP, ESR)	Secondary care	Cochrane systematic review (16 reports, 1895 children for CRP; 8 studies, 1910 children for ESR)	Summary sensitivity 0.93 (0.86 to 0.96) for CRP>20mg/L; 0.83 (0.71 to 0.91) for ESR>30mm/h. Summary specificity 0.37 (0.24 to 0.53) for CRP; 0.57 (0.41 to 0.72) for ESR. (90)
Serious infection in febrile children (CRP, ESR)	Secondary care	Systematic review (5 reports of CRP; 1379 children. 1 report of ESR)	Pooled positive likelihood ratio of raised CRP 3.2 (95% CI 2.7 to 3.7) negative likelihood ratio 0.33 (0.23 to 0.49).(91) Similar results in 2 other reviews (92, 93)
Bone and joint infections (CRP)	Secondary care	Systematic review (7 reports with 583 patients in total, of which 3 reports of CRP)	Pooled sensitivity 0.85 (0.72 to 0.94), specificity 0.37 (0.27 to 0.47). (94)
Septic arthritis (CRP, ESR)	Secondary care (emergency departments)	Systematic review (4 reports of CRP; 11 reports of ESR)	CRP sensitivity varied from 0.44 to 0.91, specificity 0.15 to 0.85; ESR sensitivity varied from 0.18 to 0.95, specificity 0.11 to 0.94, no meta-analysis performed due to heterogeneity. (95) Similar findings in another study (96)
Osteomyelitis of the foot in diabetes (CRP, ESR)	Secondary care	Systematic review (8 reports, 7 reporting ESR, 351 patients, 2 reporting CRP, 404 patients)	ESR pooled sensitivity 0.81 (0.71 to 0.88) specificity 0.90 (0.75 to 0.96) (97)
Prosthetic joint infection (CRP)	Secondary care	Systematic review (25 studies)	Pooled sensitivity 0.82 (0.80 to 0.84), specificity 0.77 (0.76 to 0.79).(98) Similar findings in a previous study which also examined ESR (99)
Chorioamnionitis in premature delivery (CRP)	Secondary care	Systematic review (23 reports; 1717 patients)	Sensitivity 0.59 (0.48 to 0.69) specificity 0.83 (0.74 to 0.89) for CRP >20mg/L (100)
Bacterial infection in patients with chronic renal insufficiency (CRP)	Secondary care	Systematic review comparing procalcitonin with CRP (7 studies; 803 patients)	Pooled sensitivity 0.78 (0.52 to 0.92) specificity 0.84 (0.52 to 0.96) (101)
Infective complications after haematopoietic stem cell transplantation (CRP)	Secondary care	Systematic review (6 reports; 1344 patients)	Pooled sensitivity 0.80 (0.54 to 0.93) specificity 0.73 (0.56 to 0.86) (102)

Bacterial endocarditis in adults (CRP)	Secondary care	Systematic review (6 reports; 1006 patients)	Pooled sensitivity 0.75 (0.62 to 0.85) specificity 0.73 (0.61 to 0.82) (103)
Excluding inflammatory bowel disease in adults with IBS (CRP, ESR)	Secondary care	Systematic review (4 reports of CRP, 823 patients; 4 reports of ESR, 616 patients)	CRP <0.5 gives a <1% probability of having IBD. No level of ESR was predictive of IBD. Sensitivity/specificity not reported (104)
Inflammatory bowel disease in adults (CRP, ESR)	Mixed primary care and secondary care	Systematic review (4 reports of CRP, 3 reports of ESR)	CRP sensitivity range 0.55 to 1.0, specificity 0.42 to 0.90. ESR sensitivity range 0.56 to 0.78, specificity 0.75 to 0.96. No meta-analysis performed due to heterogeneity. (105)
Inflammatory bowel disease in children (CRP)	Secondary care	Systematic review (9 studies; 1146 patients)	Pooled sensitivity 0.63 (0.51-0.73) specificity 0.88 (0.80-0.93).(106)
Bacterial infection in patients with systemic rheumatic diseases (CRP)	Secondary care	Systematic review (8 reports; 668 patients)	Pooled sensitivity 0.81 (0.75 to 0.86) specificity 0.63 (58.5 to 67.50) (107)
SARS-CoV-2 infection (CRP)	Secondary care	Cochrane systematic review (14 studies; 997 cases/1284 non-cases)	Pooled sensitivity 0.66 (0.55 to 0.75), specificity 0.44 (very low certainty evidence). (108)

**Where possible results are presented as pooled sensitivities and specificities to enable comparisons between studies*

1.4.3. Testing for monitoring of disease progression or treatment response

As well as diagnosis of specific diseases, inflammatory markers can be used for monitoring in patients who have inflammatory disease which has already been diagnosed. NICE guidelines suggest measuring CRP 'regularly' in people with rheumatoid arthritis, and 'monthly' for those with active disease, to inform decision-making about increasing treatment to control disease, or decreasing treatment when disease is well controlled.(109)

For infections, CRP monitoring can be used as a marker of treatment response, potentially helping to guide antibiotic therapy, particularly for infections which can be difficult to fully eradicate such as osteomyelitis.(110) Other inflammatory markers such as viscosity and ESR are less suitable for this purpose due to the long half-life and delayed response.

Some research has also evaluated the use of inflammatory markers for monitoring and assessing prognosis in diagnosed cancer.(111, 112) This tends to be done in secondary care. This thesis focuses on the diagnostic utility of inflammatory markers in primary care, so further discussion of their utility in disease monitoring is out with the scope of this research.

1.4.4. Non-specific testing for systemic disease

A third use of inflammatory markers is as a tool to differentiate between the presence or absence of disease. This can either be as a screening tool in asymptomatic patients; for example, as part of a 'health check', offered by the NHS or by private companies (the 'essential health screen' offered by citydoc.org.uk is one example). Alternatively, inflammatory markers may be used by GPs for patients with non-specific or undifferentiated symptoms as a way of trying to detect or rule-out serious underlying systemic disease. Examples of this in clinical practice include recommendations from the medical education organisation 'GP Update' to use CRP as a 'rule-out' test to exclude secondary causes of hypertension(113) or for patients with newly diagnosed memory problems prior to memory clinic referral.(114) Inflammatory marker testing is also recommended by NICE as part of the diagnostic process for chronic fatigue(115) and irritable bowel syndrome.(116) As there are no specific serological tests for these conditions, the inflammatory marker is used as a 'rule-out', with the aim of excluding other potential causes of tiredness or abdominal pains before making the diagnosis of a functional disorder.

One of the commonest non-specific symptoms seen in primary care is tiredness, and inflammatory markers are often used in primary care for these patients. Only one trial appears to have studied the appropriateness of primary care blood tests for patients with tiredness: the VAMPIRE trial. This randomised controlled trial compared immediate or postponed blood test ordering, with either limited or expanded sets of blood tests.(117) In this study the 'limited' set of blood tests

included an ESR, although it is unclear what additional diagnostic yield this test provided. Results supported restricting the number of tests ordered to this 'limited' set and have led to several guidelines recommending the use of inflammatory markers as part of the first line of investigations for patients presenting to their general practitioner with tiredness or fatigue, including the UK Clinical Knowledge Summary guidelines.(115)

Several small studies, now decades old, have evaluated the utility of inflammatory markers as a non-specific screening test. **Table 2** summarises these studies; this table was produced for my BMJ review of inflammatory marker testing(1) and updated to include a small number of more recent studies. Generally, when general practitioners test inflammatory markers for non-specific purposes the results are afterwards seen as being of little or no clinical value.(33) If the pre-test probability of disease is low then there is a higher probability that a raised result may represent a false positive. The wide range of differential diagnoses in patients without localising symptoms or signs makes interpreting 'incidental' abnormalities more difficult; this may potentially lead to 'cascade testing' (118) which may be costly or lead to patient anxiety. It is unclear based on current evidence how frequently inflammatory markers are used as a non-specific screening test in clinical practice, and what the diagnostic yield of these tests is.

Table 2: *Community and primary care studies investigating the diagnostic role of inflammatory markers as diagnostic or screening tools for non-specific disease.*

Setting	Study type	Participants	Outcome
Israeli airmen (ESR)	Prospective study, 15 year follow-up(119)	1000 healthy men aged 18-33 years: yearly ESR measurement	44 had persistently raised ESR; of these 10 subsequently developed diseases (4 myocardial infarctions, 3 ankylosing spondylitis, and one each of inflammatory bowel disease, psoriasis and benign monoclonal gammopathy)
Community study of ageing in the US (ESR)	Prospective study, 12 month follow-up(120)	100 healthy men and women aged over 70 years	9 subjects had an ESR >30 mm/h for ≥6 months; a previously undiagnosed illness was identified in 4 of these (2 polymyalgia, 1 pancytopenia, 1 anaemia)
Primary care in the Netherlands (ESR)	Prospective study(121), 3 month follow-up(122)	362 patients presenting with a new complaint for which the general practitioner considered ESR to be indicated	ESR values were on average higher in those with malignancy or inflammatory diseases. Almost all diagnoses 'revealed' by the raised ESR had been suspected at the initial consultation before the ESR result was known
Primary care in Sweden (ESR)	Prospective study(123)	439 patients over 70 presenting to primary care	58 had ESR >40 and were thoroughly investigated; of these nine new diagnoses were found and three patients could be offered effective treatment.
Primary care in Norway (ESR)	Prospective before-after study (33)	559 consultations where an ESR was performed	In 60% of the consultations the test exerted its influence mainly by supporting or reinforcing the doctor's clinical opinion. In 11% the results were unexpected and forced the doctor to reconsider. In 22% the results were felt to be of little or no clinical consequence.
Community setting in US (ESR and CRP)	Prospective study (124)	101 elderly men and women living in residential and nursing homes	CRP and ESR were significantly higher in those with infections and/or inflammatory conditions.

Despite the very limited evidence to support this practice, my pre-doctoral qualitative interviews with GPs suggest this type of testing is common in clinical practice:(2)

> *I'm fishing really. So it's, a lot of our work is early presentation of undifferentiated disease and I get, essentially buying time I get very strongly reassured, rightly or wrongly, by negative inflammatory markers.*

Clinicians talked about a fear of 'missing something', and many used inflammatory markers as a way of dealing with diagnostic uncertainty and to help reassure themselves that there was 'nothing serious going on'. Clinicians

often had the expectation that results were going to be normal and used the test to help 'rule-out' serious pathology:(2)

> *So, if we had a test that, a single blood test, that doctors could do which would reassure the patient there was nothing bloody wrong at all, then that would be a very popular test. We'll have the "nothing wrong at all" test for you, sir… You know, all the other tests are, well, you might have this specifically wrong with you or you might have this… But the CRP is probably the closest thing that we've got to a "nothing wrong at all" test.*

1.4.5. Inflammatory markers as a predictor of future disease

Another area of research interest is the utility of inflammatory markers as a predictor of the risk of future disease, including cardiovascular disease,(125) diabetes,(126) atrial fibrillation,(127) cancers,(128-130) dementia, (131) schizophrenia,(132) depression(133) and overall mortality.(134) At present this is mostly of research interest and has not yet entered mainstream clinical practice, however it has potential to lead to significant changes in testing rates in future.

In primary care, the main potential areas of clinical interest are the identification of patients at increased risk of mortality and those at increased risk of hospital admission, particularly given the context of an ageing population and rising multimorbidity. Predictive tools could help inform the planning and delivery of services including advanced care planning, and interventions to reduce hospital admissions. An inflammatory marker test (ESR) is included as part of QAdmissions, a tool for predicting hospital admissions in primary care populations.(135) Several mortality risk tools exist: none in current use includes an inflammatory marker test.(136, 137) There is evidence for an association between inflammatory markers and mortality in secondary care,(138) and evidence of a small association over the long term from population based cohort studies.(139) The association in a primary care population, and over the shorter term, is unknown.

The role of inflammation in the susceptibility and pathogenesis of cardiovascular disease has also received much attention. A wide variety of inflammatory biomarkers have been studied, and high-sensitivity C-reactive protein (hs-CRP) has been examined for its potential role in risk stratification. The hs-CRP test measures the same plasma protein as the 'standard' CRP test, but has a lower limit of detection for the assay. A number of professional societies have produced guidelines or recommendations for hs-CRP testing for cardiovascular risk profiling. (140, 141) This has not yet entered mainstream clinical practice in the UK.

If hs-CRP is incorporated into mainstream clinical practice for cardiovascular risk profiling it will lead to increased testing of asymptomatic patients; inevitably a proportion of them will be found to have unexpected abnormalities. When results are in the peri-normal range, they can be used for cardiovascular risk profiling; however there has been a lack of attention paid to the issue of how to best manage the patients who are found incidentally to have raised CPR results. In asymptomatic patients with a low pre-test probability of disease, there is a higher probability that abnormal results represent false positives. Present guidelines offer relatively little advice for clinicians, for example; '*because hs-CRP can be elevated during acute illness, clinical judgment should be exercised in the interpretation of any single measurement of hs-CRP*'(140) and '*further assessment of patients with highly elevated hs-CRP (>10 mg/L) for non-cardiovascular causes of inflammation was also endorsed*.'(142)

Further research to determine an optimal evidence-based approach to the interpretation and management of raised inflammatory markers is therefore timely. However, as this thesis focuses on the diagnostic utility of raised inflammatory markers, further discussion of their role in cardiovascular risk stratification will not be explored.

1.4.6. Comparative diagnostic accuracy of inflammatory markers

Comparison of the accuracy of CRP versus ESR was examined in a recent systematic review.(143) A total of 29 studies were identified; 16 measuring accuracy of ESR versus CRP for orthopaedic infections, 3 for rheumatic disease and 10 'other' diseases; all studies were based in secondary care settings. CRP did not appear to have a superior diagnostic accuracy to ESR for orthopaedic infections, with area under summary receiver operating curve (SAUC) 0.80 for ESR versus 0.81 for CRP. Despite high heterogeneity in the 'other' category (which included pyelonephritis, pneumonia, sepsis, meningitis and epididymitis), meta-analysis for this group was performed, with higher SAUC for CRP (0.86) compared with ESR (0.75) leading the authors to conclude, on limited evidence, that CRP had superior diagnostic accuracy.

Another recent comprehensive systematic review aimed to determine when it is appropriate to concurrently test ESR and CRP (as opposed to testing only ESR or CRP) to help diagnose inflammatory disease or serious infection.(144) Evidence for single versus multiple inflammatory marker tests was identified for a limited number of conditions; suspected periprosthetic joint infections, paediatric orthopaedic infections, giant cell arteritis and inflammatory bowel disease. Most of the evidence was from secondary care and was assessed as being at high risk of bias. Due to the limited number of direct comparisons between ESR and CRP which could be made within studies, the authors also made indirect comparisons of the accuracy of CRP and ESR across studies with the same disease condition. As a result of the limited evidence identified the authors were unable to make any clear recommendations about preferred choice of tests.(144)

Limited published evidence comparing the diagnostic accuracy of PV versus CRP and ESR is available, including a small study on periprosthetic joint infections,(145) a case-control study on multiple myeloma(59) and a systematic review (without meta-analysis) on inflammatory arthritis.(146)

In summary there is very limited evidence comparing the diagnostic accuracy of inflammatory markers in primary care, and a lack of evidence to determine the added benefit of testing multiple inflammatory markers simultaneously. Due to the lack of guidelines and uncertainty regarding which inflammatory marker to use, clinicians interviewed in my pre-doctoral qualitative research reported frequently using two inflammatory markers simultaneously:(2)

> *I probably tend to do both [CRP and PV], actually, which is probably more out of just making sure I caught everything.*

1.4.7. Interpretation of inflammatory marker tests results

The interpretation of a raised inflammatory marker should be relatively straightforward if there is a clear pre-test hypothesis against which the test result can be evaluated – for example assessing the likelihood of pneumonia in a patient with a cough. This was demonstrated by a Dutch study of patients in whom the raised ESR seemed to confirm an initial diagnosis as opposed to revealing unexpected disease.(121, 122) The difficulty lies in interpreting an 'incidental' finding, when no specific disease is suspected. Markedly raised inflammatory markers (such as an ESR >100mm/h or CRP>100mg/L) have been shown to have a high likelihood of disease. The conditions found depend on the setting, but include infection (33-60%), inflammatory disease (14-30%) and malignancy (5-28%)(147-152). In my pre-doctoral qualitative research, GPs I interviewed identified interpretation of a raised inflammatory marker in primary care as a particular challenge, which could lead to lead to uncertainty, follow up GP consultations and further testing:(2)

> *'Then you think suddenly, well should I be looking further and further and further, but that could mean more and more random investigations until you get the point where you goes, oh, I'll just do a whole body CT scan to see if anything pops up I suppose.'*

1.5. Qualitative research and inflammatory marker testing

Given the large numbers of quantitative research studies exploring inflammatory marker testing summarised above, it is perhaps surprising that relatively little qualitative research has been done. To address this lack of qualitative research into inflammatory marker testing I conducted interviews with GPs about these tests in my pre-doctoral Academic Clinical Fellowship.(2)

Further qualitative research to address both doctors' and patients' perspectives is important for several reasons. Firstly, qualitative research can be used to explore the psychosocial reasons driving use of tests from both doctors' and patients' perspective. This is important because understanding the diagnostic accuracy of a test is necessary but not sufficient; to improve test usage non-medical drivers for testing must be understood and addressed. This was summarised in 1998 by Little *et al* who said: '*If psycho-social agendas are important in ordering investigations, then clinical guidelines which discuss only medical criteria may not be effective in reducing 'inappropriate' investigations*'. (153) Despite this, the vast majority of guidelines and interventions to 'optimise' use of blood tests still focus on medical reasons for testing.(154)

Secondly, qualitative research is important because tests in themselves do not make people better - unless actions based on the test result lead to either a change in patient management or reassurance. Both of these are dependent on test result communication. A systematic review of US studies quantifying failures in test result follow up has shown that between 6.8% and 62% of laboratory tests are not followed-up; no relevant UK research was identified.(155) Surveys and qualitative studies have shown that UK general practices generally rely on patients contacting the practice to obtain their test result, with a lack of fail-safe mechanisms.(156-158) Errors associated with filing, communicating and actioning of abnormal results could lead to delayed and missed diagnoses;(159) conversely if normal results are not adequately

communicated patients are unlikely to be reassured by testing. Improving the use of tests therefore requires improved communication around testing.

Finally, shared decision-making has been described as 'the pinnacle of patient-centred care',(160) yet most research into shared decision-making focuses on treatment decisions rather than diagnostic testing.(161) If shared decision-making between clinician and patient is to be achieved it is crucial that doctors and patients have a shared understanding of why inflammatory marker tests are being done and the meaning of the results.

1.5.1. Doctors' perspectives of blood tests

Studies exploring doctors' reasons for testing have been reviewed by Whiting *et al* and classified into several key groupings: diagnostic factors, therapeutic and prognostic factors, patient-related factors, doctor-related factors and policy and organisation related factors.(162) From this it is clear that diagnostic factors, which can be modelled mathematically in terms of Bayesian pre-test and post-test probabilities, are only one factor in a complex decision-making process. A telephone survey of 300 physicians in the US indicated that most do not use formal recommended quantitative methods for appraising a diagnostic test, with only 3% reporting that they use formal Bayesian calculations.(163) In my pre-doctoral qualitative work I explored these non-medical motives for testing by looking at the question '*what do tests do for doctors?*'.(164) I found that doctors viewed tests as a way to manage uncertainty within the context of increased litigation, risk aversion and reduced continuity of care. They could also be used as part of the social interaction as a 'gift' for the patient, and a way to be seen to be 'doing something', within the social context of time pressures and perceived patient pressures. A recent realist review by Duddy *et al* explored clinicians' decision-making in relation to laboratory testing, and found that workload and time pressures tend to promote the use of tests, with clinicians prioritising efficiency over thoroughness in decision-making.(165)

1.5.2. **Patients' perspectives of blood tests**

There are a smaller number of studies looking at patients' perceptions of blood testing, with none specifically focussed on inflammatory marker testing. One qualitative study interviewed patients recruited via GP practice waiting rooms, and found that patients tended to overestimate the qualities of blood tests, which were seen as almost 'magic' and regarded as providing diagnostic certainty.(166) This is in keeping with a systematic review of quantitative studies measuring patient expectations of tests and treatments; the majority showed patients tend to overestimate benefits and underestimate harms.(167) Studies exploring whether patients are reassured by normal test results have shown some evidence of short term reductions in anxiety and GP visits, for example with neuroimaging in chronic daily headaches.(168) However, systematic reviews have shown that overall normal diagnostic tests do little to reassure patients, decrease anxiety or resolve their symptoms.(169, 170)

Comparisons between doctors' and patients' perspectives are important to highlight potential areas of misunderstanding, which can impact on decision-making. Qualitative studies into antibiotic prescribing have shown that patients and doctors can sometimes be talking at cross purposes about the 'seriousness' of the illness, with patients' expressions of anxiety misinterpreted as a pressure for antibiotics.(171) Similarly, one study measured patients' expectations in general practice and found that patients have less desire for 'tests and diagnosis' and a greater desire for 'explanation of the problem'.(172) It is therefore possible that patients' expressions of anxiety may be misinterpreted as a pressure for tests.

1.6. Shared decision-making and informed consent

The General Medical Council published guidelines in 2020 for decision-making and informed consent which recommend that doctors and patients should *'reach a shared understanding of the expectations and limitations of the available options'*.(173) The legal framework underpinning these guidelines comes from Montgomery v Lanarkshire (2015), which state that '*consent must be obtained before treatment interfering with bodily integrity is undertaken'*, with doctors obliged to ensure patients are aware of risks and benefits of each course of action that are meaningful to the particular person.(174) The amount of information sharing required for blood tests is unclear, although there have been cases of complaints and litigation in primary care, for example, patient who was not informed that they were having liver function tests done alongside cholesterol screening bloods.(175)

Shared decision-making goes above and beyond this legal framework for informed consent and shared understanding, seeking not only to inform but to actively involve and engage patients in decision-making. Shared decision-making has been defined as '*a collaborative process that involves a person and their healthcare professional working together to reach a joint decision about care'*.(176) Evidence for the benefits of shared decision-making is mostly based around treatment decisions; (161) however, draft NICE guidelines are not limited to treatment decisions and recommend shared decision-making '*to choose tests, treatments, management or support packages, based on evidence and informed personal preferences, health beliefs and values*'.(176)

Evidence for shared decision-making around blood testing is limited, although specific tests such as prostate specific antigen (PSA),(177) genetic testing(178) and screening tests(179) have received much attention, and some decision aids are available.(180) It is unclear whether the same principles are applicable to blood tests such as inflammatory marker tests. Observational research using video-

recorded GP consultations in UK primary care has demonstrated that the reality of decision-making around blood testing does not reflect the ideals of shared decision-making, with most decisions being doctor-led, with a lack of information-sharing and shared decision-making.(181, 182)

1.7. Gaps in current evidence base

Despite the large numbers of studies evaluating inflammatory markers summarised in section 1.4, there is a lack of clinically relevant research for general practice. Most research on inflammatory markers comes from a secondary care setting, where disease prevalence is higher. This evidence cannot be extrapolated to a low-prevalence primary care setting, due to the variation in test performance seen amongst different population subgroups, which is known as the spectrum effect.(183) The impact of this spectrum effect is that tests developed in high prevalence settings (e.g. secondary care) will typically have a lower sensitivity and higher specificity when applied in a population with lower disease prevalence (e.g. primary care).(183)

Additionally, most studies measure accuracy for single disease outcomes; however, this does not reflect the clinical experience of general practitioners who are faced with a test result and often need to weigh up the risks of a wide range of possible disease outcomes, to decide on the most appropriate further tests and treatments required. In clinical practice this leads to uncertainty and anxiety for clinicians. They describe a tension between not wanting to 'miss anything', and, on the other hand, being wary of picking up borderline abnormalities which can lead to cascades of further tests.(2) Whilst clinicians commonly use inflammatory markers as a 'rule-out' or 'nothing wrong at all' test, there is a lack of evidence to support this practice. There is therefore uncertainty within the current research literature about when to use inflammatory markers and how to interpret results.

Another limitation of the current inflammatory marker literature is that diagnostic test accuracy (DTA) studies generally present results in terms of sensitivity and specificity, which have been shown to be poorly understood by clinicians and patients, are difficult to interpret and rarely used in clinical practice.(163) Research suggests that natural frequencies are easier to understand than sensitivities and specificities, and methods of presenting DTA studies using natural frequencies have been developed.(184)

Additionally, current DTA study methodologies use a binary model of test positive versus test negative. This ignores those with indeterminate test results which can be the most challenging to manage in primary care, and may lead to overestimation of the probability of disease in patients who have borderline abnormal results, and underestimation in those with significant abnormalities. My qualitative research highlighted that this was perceived as a problem with inflammatory marker testing:(2)

> *So, I think that if they're very normal that's good, if they're very abnormal that's helpful. I think that probably where the problem lies is if you get like a mildly raised viscosity and you wonder is that, how significant is that?*

Finally, there is a lack of evidence addressing patients' perspectives of testing. Although shared decision-making is increasingly promoted as best practice, there is also a lack of research exploring shared decision-making in the context of diagnostic decisions.

In this thesis I set out to address some of these limitations. I use routine data from UK primary care to explore how tests are used in clinical practice, and to measure the diagnostic accuracy for a combined disease outcome of any infection, autoimmune disease or cancer ('relevant disease'). The results are presented in a variety of ways including natural frequencies to aid understanding and applicability. As well as disease outcomes, I measure process outcomes including rates of consultations, blood test and referral after inflammatory marker testing to quantify the cascade effect. Additionally, as well as looking at a binary model of test positive versus test negative, I also explore

the dose-response relationship between test results as a continuous variable and disease incidence. Finally, by incorporating qualitative methods I explore patients' and doctors' perspectives and generate finding to help doctors and patients develop a shared understanding of testing.

1.8. Ontological and epistemological position

Ontological and epistemological beliefs influence how research questions and methods are shaped and interpreted. Both are philosophical concepts – ontology relates to our knowledge and understanding of the world; epistemology relates to the methods, validity and scope of our knowledge. My thesis incorporates quantitative and qualitative methods which come from different epistemological positions. Quantitative research is generally grounded in the positivist paradigm, which recognizes only that which can be scientifically measured and verified with empirical data or which is capable of logical or mathematical proof. This is in keeping with my background as a doctor and my training in evidence-based medicine and the biomedical model of disease. The focus of the quantitative component of my thesis, to measure the diagnostic accuracy of inflammatory marker tests, is in keeping with this positivist paradigm.

The qualitative component of my thesis however aligns more with an interpretivist paradigm, in which the way the world is experienced and understood is socially constructed and depends on how people interpret and make sense of their experiences.(185) This is in keeping with my training and experience as a general practitioner, in which patients' experience of illness is not solely explained by biomedical pathology, but influenced by their beliefs, experiences, social interactions and psychology.

My overall epistemological approach is therefore one of methodological pragmatism:(186) I believe that in order to understand the clinical practice of inflammatory marker testing, both qualitative and quantitative methods are

needed. My priority has been to ensure that the research questions and the results of my research are relevant to clinicians and patients. I believe a pragmatic approach is needed to choose optimal methods to achieve this, even though this means being flexible and willing to switch between paradigms. I feel comfortable switching between positivist and interpretivist paradigms as a GP, as this reflects my daily clinical practice of medicine. In my clinical practice I draw on my biomedical knowledge, from a positivist perspective, to interpret symptoms, signs and test results and to recommend evidence-based treatments. At the same time, I also explore patients' ideas, concerns and expectations, which brings an interpretivist approach to understanding and exploring their problems. My clinical experience has taught me that a biomedical understanding of disease is necessary but not sufficient. Holistic care of patients requires integration of biomedical knowledge with psychosocial factors which influence patients' experience of illness. In the same way, my research experience has shown me that different types of knowledge, from both positivist and interpretivist paradigms, bring important insights which can be used to shed light on research questions.

1.9. Aims and Objectives

1.9.1. Overall aims

The overall aim of this thesis is to explore the diagnostic utility and clinical practice of inflammatory marker testing in primary care, including how results are shared with patients, using a mixed methods approach. This will inform policies and clinical practice to optimise the use of inflammatory marker tests in primary care.

1.9.2. Quantitative aims and objectives

For the quantitative component of my thesis, I use a cohort of 160,000 patients from the Clinical Practice Research Datalink (CPRD) who had inflammatory marker blood tests in 2014, and 40,000 matched untested patients. The overall aim is to determine the diagnostic utility of inflammatory marker tests in primary care for relevant disease, including infections, autoimmune disease and cancer. The results will be of relevance for a general practitioner deciding when to use inflammatory markers and how to interpret results.

Specific objectives for the quantitative phase of the thesis are:

1) To describe the baseline characteristics in terms of age, gender, socioeconomic status and ethnicity of patients having inflammatory marker tests in primary care.
2) To determine the diagnostic accuracy of CRP, ESR and PV, singly and in combination, for relevant disease (infections, autoimmune conditions or cancer) in primary care, and to compare disease incidence in tested versus untested populations.
3) To determine the symptomatology of patients with inflammatory marker testing in primary care and measure the consequences of testing in terms of numbers of consultations, blood tests and referrals.
4) To determine the diagnostic accuracy of inflammatory markers specifically for cancer in primary care, including stratification by age, gender, inflammatory marker level and cancer type.
5) To explore the association between inflammatory markers and one-year mortality in primary care.

1.9.3. Qualitative aims and objectives

The second component of the thesis is a qualitative study, involving interviews with patients who had recent inflammatory marker tests at two time points: (a) at or soon after their blood tests and (b) after they received their test results. I also interview the GPs who requested these tests. The overall aim is to explore the meaning of inflammatory marker tests for doctors and patients in primary care.

The specific objectives of this qualitative component are:

1) To explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests.
2) To provide in-depth exploration of patients' experience of testing - from GP consultation to results.
3) To identify barriers and facilitators to communication and shared understanding, in order to inform improved communication in future.

1.10. Outline of chapters

In the next chapter I will describe the methods of the quantitative component of my thesis, in chapter three I will then describe the results. Chapter four describes the methods of the qualitative component of the thesis and chapter five describes the results. Finally, in chapter six I will discuss and synthesise the findings of both the qualitative and quantitative components of my thesis, considering the strengths and limitations of the methods chosen, comparisons with existing literature and the implications for clinical practice and future research.

CHAPTER 2. QUANTITATIVE METHODS

2.1. Chapter overview

This chapter describes the methods used for the quantitative study. The overarching aim was to determine the diagnostic utility of inflammatory markers in primary care populations for relevant disease, including infections, autoimmune disease and cancer. The specific objectives are described in section **1.9.2**.

To achieve these objectives, I used a cohort of 160,000 patients from the Clinical Practice Research Datalink (CPRD) who had an inflammatory marker blood test in 2014, and 40,000 matched untested patients, and measured the incidence of relevant disease following testing.

In this chapter I first provide an overview of the study design and sources of data. Then I describe the methods of code-list development, stages of data manipulation, variables created, and finally the statistical methods and analyses used.

2.2. Study design

This was a diagnostic cohort study to measure the accuracy of inflammatory marker tests in primary care using electronic health record data from CPRD. The index test was defined as any inflammatory marker blood test (CRP, ESR or PV) performed in 2014. The target condition was any relevant disease defined as diagnosis of one or more of inflammatory conditions including infection, autoimmune disease, or cancer. The reference standard for diagnostic accuracy

research is usually defined as the best available method for classifying whether people have the target condition or not. In this study one-year incidence of cancer and autoimmune disease and one-month incidence of infection was used as a proxy measure for presence of disease at the time of testing; thus, the reference standard is based on GP diagnosis and coding of disease. This was chosen as a reasonable compromise between identifying all the conditions related to the raised inflammatory marker yet reducing the chance of identifying unrelated conditions.

2.2.1. **Sources of data**

The CPRD is a government run database containing anonymised routinely collected data from primary care electronic health records. It contains coded information about patient demographics, diagnoses, symptoms, prescriptions, referrals, and test results. It also offers linkages to data sources outside primary care including to data recorded by the Office for National Statistics (ONS), the National Cancer Registration Service (NCRS) and Hospital Episode Statistics (HES), although not all patients are linked to all datasets.

At the time of analysis, CPRD GOLD included over 11.3 million patients from GP practices using Vision® software, of whom 4.4 million were active (alive and currently registered) and meeting quality standards. This comprised approximately 6.9% of the UK population, and has been shown to be representative of the UK general population in terms of gender, age and ethnicity.(187) More recently CPRD Aurum has been developed, containing data from practices using EMIS Web® software; this was not available at the time of analysis so this study is limited to CPRD GOLD.(188)

2.2.2. Study population

Eligibility criteria for the study were as follows:

- Aged 18 or over in 2014.
- Defined as 'up to standard' according to CPRD criteria.
- Had inflammatory marker blood test (CRP, ESR or PV) performed between 01/01/2014 and 31/12/2014 (identified using the code-list in Appendix A).

At the time of extraction (23rd June 2017), 463,304 patients fulfilled these eligibility criteria, of whom 160,000 patients were chosen at random by CPRD. A comparison sample of 40,000 untested patients were selected by CPRD – these were patients who had no inflammatory marker test in 2014, although they could have had testing at other dates. These were matched by age (in 5-year bands), sex and practice, to a random subset of 40,000 patients from the inflammatory marker test group. No exclusion criteria were applied, in order to maximise generalisability.

2.2.3. Rationale for the untested comparison group

The untested comparison group allowed me to quantify the disease risk associated with inflammatory marker testing, as well as the risk associated with test results. This is important as diagnostic testing is a two-step Bayesian process. The first step is the clinicians' decision to perform an inflammatory marker test on a patient, based on the pattern of symptoms, signs and risk factors in an individual, which clinicians use to generate an overall 'gestalt' judgement. The second step is the test result which is used to refine this judgement. Comparing the incidence of disease in the tested versus untested populations, gives an indication of the diagnostic accuracy of clinicians' gestalt judgement. If the excess disease risk associated with testing is not fully eliminated by a negative

result, those with normal inflammatory marker test results could have increased disease risk compared to untested. This has been demonstrated with other tests, for example male primary care patients with a normal platelet count (<400x109/l) have a one-year cancer risk of 4.1%; above the NICE threshold for urgent referral.(189) This is particularly important, as clinicians tend to use inflammatory markers as a 'rule-out' test.

Matching was required for this comparison group as we know that age and sex are strong determinants of disease, and because different practices have different recording and testing styles. The alternative of requesting a comparison group never to have had an inflammatory marker test would introduce a bias by including patients with particularly good health, as nearly 25% of potentially eligible patients aged >18 in CPRD have had an inflammatory marker test at some time.

2.2.4. Index tests, target conditions and co-variables

The index tests were inflammatory marker blood tests: CRP, ESR and PV. The index date was defined as the date of the first inflammatory marker blood test performed in 2014. Controls were allocated the same index date as their matched case.

The overall target condition was 'any relevant disease', defined as any cancer or autoimmune conditions coded within 1 year, or infection within 1 month of the index date. A secondary outcome of cancer incidence at 2-years after inflammatory marker testing was measured, to detect any late effect. As well as examining the overall diagnostic accuracy for 'any relevant disease' I also calculated measures of diagnostic accuracy for infections, autoimmune conditions and cancers separately.

Age, gender and socioeconomic status were examined as co-variates. I also analysed symptomatology at the time of testing by measuring the top 20

symptoms in the 28 days before testing in the tested and untested groups, to give an indication of the reasons GPs are ordering the tests. Process outcomes of repeat GP consultations, additional blood tests and referrals, were identified, in order to quantify the cascade effects of testing.

2.2.5. Sample size calculation

I performed a sample size calculation based on estimating the 12-month condition incidence amongst those with raised inflammatory marker. CPRD feasibility data indicated that roughly one third of inflammatory marker tests in 2014 were above the laboratory normal range. Using this assumption of a 2:1 ratio of test negative: test positive, I calculated that a total sample size of 77,577 would achieve an 80% power to detect a 0.2% increase in the incidence of a condition in the test positive compared with the test negative group, assuming a baseline (test negative group) condition incidence of 0.8% and with a type II error rate of 5%. This is realistic given that previous studies have shown that the positive predictive value of a raised inflammatory marker for myeloma is 0.2 (Ref: NICE Cancer guidelines); for multiple cancers I would expect this to be higher. I rounded this to a sample size of 80,000 to allow for dropouts once those with pre-existing relevant disease were excluded from analyses.

As well as examining the three inflammatory markers combined, I also wanted to look at diagnostic accuracy of each of the three tests individually. Of the three, plasma viscosity was the least used, comprising 8% of the total tests, or 6,400 out of the 80,000 total sample size. After exclusions I would expect to have over 4977 plasma viscosity tests. This sample size achieves an 80% power to detect (with a type II error rate of 5%) a difference of 2% disease incidence in the test positive group, versus 1% in the test negative group; ample power for the more common outcomes.

A comparison group of 20,000 matched untested patients was requested from CPRD. I opted for a smaller sample size in this group to ensure maximum power

in the main study, as this comparator group was not used for the majority of the analyses, and no subgroup analyses were done on this group. Based on the assumption of a condition incidence of 0.8% in the inflammatory marker test group (as above), in order to detect a condition incidence in the untested group of 0.6% with a 4:1 ratio (allowing for the proposed smaller sample size of the comparison group) requires 65,772 in the inflammatory marker tested group, and 16,443 in the untested comparator group for an 80% power; rounded to 20,000 to accommodate exclusions.

Following feedback from the CPRD the sample size was increased to 200,000 (160,000 tested and 40,000 untested cohorts). They suggested that this was within the same costing envelope and would increase the power to detect rarer outcomes such as cancers and autoimmune conditions.

Sensitivity analysis to check sample size

As my primary objective was to measure the diagnostic accuracy of inflammatory marker tests, I did additional sensitivity analyses to check my proposed sample size (**Table 3**). I used the following formula to check the likely standard error for my proposed sample size under a variety of assumptions for prevalence and sensitivities:

$SE = \sqrt{(p(1-p) / k)}$

k = expected number of cases (sample size x prevalence)

p = sensitivity

SE = standard error

Table 3: *Modelling for standard error given a sample size of 160,000*

		Prevalence					
		0.25%	**0.5%**	**1%**	**2%**	**5%**	**10%**
Sensitivity	**50%**	2.5%	1.8%	1.3%	0.9%	0.6%	0.4%
	60%	2.5%	1.7%	1.2%	0.9%	0.5%	0.4%
	70%	2.3%	1.6%	1.1%	0.8%	0.5%	0.3%
	80%	2.0%	1.4%	1.0%	0.7%	0.4%	0.3%
	90%	1.5%	1.1%	0.75%	0.5%	0.3%	0.2%

This demonstrates that my sample should give sufficient power to detect rare outcomes such as cancer and autoimmune conditions down to prevalence rates <1% with test sensitivities of up to 90%.

2.2.6. **Ethics approvals**

Approval for this study was obtained from the Independent Scientific Advisory Committee (ISAC) approval reference number 17_003.

2.2.7. **Reporting**

The reporting of this study conforms to the STARD(190) and RECORD(191) statements.

2.3. Code-list development

The methods of code-list development described below have been published in the following Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license:

- Watson J, Nicholson BD, Hamilton W, & Price S. (2017) Identifying clinical features in primary care electronic health record studies: methods for codelist development. *BMJ Open* 7(11) https://doi.org/10.1136/bmjopen-2017-019637

Code-list development is an important first step for most research using electronic health records. This apparently simple task is made complex by the fact that methods for developing code-lists are not standardised and are often poorly reported. In order to improve transparency, replicability and rigor in code-list development I published a standardised method for code-list development, based on methods used by my supervisor Professor Willie Hamilton and his team.(192) This method involves three stages, summarised in **Figure 1**.

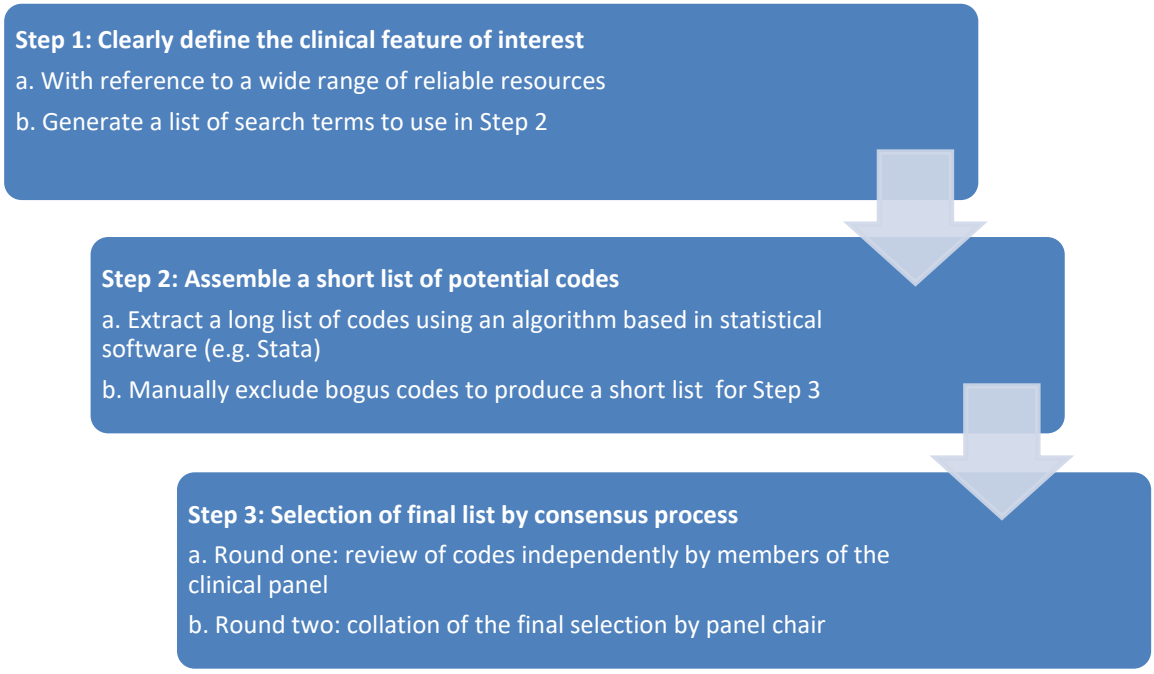


Figure 1: *The method for code-list collation consists of three steps*

2.3.1. Step 1: clearly defining the target condition

To identify the target condition, the first stage was to collate lists of codes pertaining to infections, autoimmune conditions, and cancers. To achieve this, I first needed to clearly define the diseases of interest and inclusion and exclusion criteria. **Appendix A** contains a summary of the disease outcomes for infections and autoimmune conditions which were studied. Pre-existing cancer code-lists were available (see **2.3.4**) so this process was not repeated for cancers.

Reliable sources of clinical information were used to define each disease outcome including:

- International Classification of Primary Care (ICPC)
- National Institute for Health and Care Excellence (NICE) Clinical Knowledge Summaries (CKS)
- ICD-10
- GP notebook

2.3.2. **Step 2: assemble a list of potential codes**

The second stage consisted of identifying all potential codes that might be used by GPs to record the target condition defined in Step 1 and collating them into a list. This was done using Stata to search the lookup file (medical.txt) provided by CPRD. This contains the alphanumeric 'Read code' which is used in GP electronic health records for storing coded information, the CPRD's 'medcode' (which is simply a numeric equivalent of the Read code) and a text description ('desc') which is common to both the readcode and the medcode.

Using the resources listed in Step 1, I developed an exhaustive list of synonyms for each disease of interest. Using Stata, I created a variable for the disease outcome of interest and set this to zero. I then used Stata to search the description of each code, and set this new variable to 1 if it contains any of the synonyms. **Figure 2** gives example syntax used for systemic lupus erythematosus (SLE).

```
import delimited “medical.txt", clear

*generate a binary variable for SLE and set its value to zero
g autoimmune_SLE=0

*search the description of the medcode and change the value of variable
autoimmune_SLE from 0 to 1 if it contains words that suggest the code might be
related to the diagnosis of SLE

replace autoimmune_SLE=1 if regexm(desc, "SLE")
replace autoimmune_SLE=1 if regexm(desc, "[Ll]upus|LUPUS")
replace autoimmune_SLE=1 if regexm(desc, "[Ll]ibman|LIBMAN")
*order the dataset so that values of variable autoimmune_SLE==1 are all placed
together
gsort autoimmune_SLE

*manual check for bogus codes – manually change to autoimmune_SLE=0 if the code
is clearly inappropriate

*drop bogus codes
drop if autoimmune_SLE==0

*retain the variables of interest
Keep medcode readcode desc autoimmune_SLE

*save the file as a library
save  "Autoimmune_SLE.dta", replace

*export as an excel file for step 3 (see below)
Export excel using “autoimmune_SLE”, replace
```

***Figure 2**: Example syntax used to identify medcodes for SLE*

To cross reference and validate this, the same Stata do-file was used to search an additional file which contained ICD-10 codes cross-referenced to Read code and medcodes. As well as the text ‘description’ this file also contained a variable ‘ICD description’; both were searched for synonyms representing the disease of interest. Additionally, the code-list was then sorted by ICD-10 codes, which allowed relevant sections of ICD-10 to be visually inspected to ensure that no additional missing codes had been omitted. In some cases, this cross-check

helped to generate additional synonyms for searching in an iterative process. Unfortunately, this file is no longer provided by CPRD so was not fully up to date; however, using the latest CPRD lookup file ('medical.txt') ensured that newer codes were not omitted.

Once a complete list of potential codes had been produced this was manually inspected for any bogus codes which were clearly inappropriate. This manual check for bogus codes erred on the side of caution, only rejecting codes that were clearly inappropriate according to predefined inclusion and exclusion criteria. Common reasons for exclusion were codes indicating a family history of a condition, absence of a condition, or screening for a condition rather than presence of a condition.

The output from Step 2 was then exported to Excel as a list of potential codes for clinical review (Step 3).

2.3.3. Step 3: selection of final list by consensus process

Each code-list was reviewed by me, plus at least one other GP from a panel of six using a modified nominal group technique.(193) Each GP independently categorised the list, ranking each Read code/medcode using a 3-point scale as follows:

1 = Definitely include - the code accurately defines the clinical feature of interest, and GPs would definitely use it.
2 = Uncertain – it remains unclear whether the code accurately reflects the clinical feature of interest, or whether GPs would use it.
3 = Definitely exclude – the code does not define the clinical feature of interest, and GPs definitely would not use it.

Panel members were encouraged to add comments to explain the reasons for their inclusion/exclusion. This was shared with the panel chair and helped in cases of uncertainty to clarify whether disagreements reflected misunderstandings or problems with definition or inclusion/exclusion criteria,

or whether the uncertainty reflected the inherent variability in clinical coding between clinicians.

Codes were retained in the final list if they were ranked as '1=definitely include' by both reviewers. Codes ranked as '3=definitely exclude' by both reviewers were dropped. Remaining codes were discussed with my supervisor Professor Willie Hamilton, a GP with significant experience of code-list collation, and final decisions regarding inclusion/exclusion was made by consensus.

Full code-lists for infections and autoimmune conditions are published on the University of Bristol Research Data storage repository.(194) **Table 4** shows the complete results of the consensus review for the example of SLE.

Table 4: *List of potential codes for inclusion to identify patients with SLE*

(1: definitely include, 2: uncertain, 3: definitely exclude)

medcode	Description	1st reviewer decision	2nd reviewer decision	Final decision	Comments
2667	Discoid lupus erythematosus	1	3	1	As discoid lupus is autoimmune, we clarified the definition of the disease outcome to include this
4125	Lupus erythematosus	1	1	1	
7522	Lupus erythematosus NOS	1	1	1	
7871	Systemic lupus erythematosus	1	1	1	
11920	Systemic lupus erythematosus with pericarditis	1	1	1	
11920	Systemic lupus erythematosus with pericarditis	1	1	1	
16367	Tuberculosis - lupus vulgaris	2	3	0	Included in the infection code-list rather than autoimmune
20007	Disseminated lupus erythematosus	1	1	1	
22205	Lupus nephritis	1	1	1	
22205	Lupus nephritis	1	1	1	
25390	Subacute cutaneous lupus	1	1	1	

	erythematosus				
28952	Lupus pernio	2	3	0	Included in the infection code-list
29519	Systemic lupus erythematosus with organ or sys involv	1	1	1	
31564	Lung disease with systemic lupus erythematosus	1	1	1	
31564	Lung disease with systemic lupus erythematosus	1	1	1	
33449	Lupus erythematosus chronicus	1	1	1	
36942	Drug-induced systemic lupus erythematosus	1	1	1	
37492	Tuberculosis - lupus NOS	1	3	0	Included in infections
40797	Lupus erythematosus migrans	1	1	1	
41148	Renal tubulo-interstitial disorder in SLE	1	1	1	
41148	Renal tubulo-interstitial disorder in SLE	1	1	1	
42719	Systemic lupus erythematosus NOS	1	1	1	
44095	Polyneuropathy in disseminated lupus erythematosus	1	1	1	
44095	Polyneuropathy in disseminated lupus erythematosus	1	1	1	
44984	Lupus erythematosus tumidus	1	2	1	
45726	Systemic lupus erythematosus disease activity index	1	1	1	
46148	Lupus erythematosus profundus	1	3	1	
46675	Lupus - tuberculous	1	3	0	Infection
47047	BILAG - British Isles lupus assessment group score	1	1	1	
47672	Nephrotic syndrome in systemic lupus erythematosus	1	1	1	
47672	Nephrotic syndrome in systemic lupus erythematosus	1	1	1	
51798	Systemic lupus activity measure	1	1	1	
57675	Libman-Sacks disease	1	1	1	
57675	Libman-Sacks disease	1	1	1	
58706	[X]Other forms of systemic	1	1	1	

	lupus erythematosus				
63283	SLEDAI-Sys lup ery dis act ind	1	1	1	
63955	Lupus erythematosus unguium mutilans	1	1	1	
65391	Lupus erythematosus nodularis	1	1	1	
67637	Tuberculosis - lupus exedens	1	3	0	Infection
94751	Eyelid discoid lupus erythematosus	1	3	1	Discoid lupus included as autoimmune – see above
103784	Lupus insensitive activated partial thromboplastin time	3	3	0	
14471	Lupus circulating anticoagulant index	3	3	0	
14478	Lupus anticoagulant screen	3	3	0	
107468	Lupus anticoagulant negative	3	3	0	
107447	Lupus anticoagulant positive	3	3	0	
38264	Lupus inhibitor activity	3	3	0	
30919	Lupus anticoagulant screening test	3	3	0	
99435	Neonatal lupus erythematosus	3	3	0	
101433	Cerebral lupus	1	1	1	
106086	SLAM - Systemic lupus activity measure	1	1	1	

2.3.4. **Cancer code-list**

A pre-existing, validated list of 2,120 cancer-related medcodes was used to identify cancer diagnoses. These were developed by my supervisor Professor Willie Hamilton using the methods outlined in section **2.3** above for the Diagnosis of Symptomatic Cancer Optimally (DISCO) studies,(195) and updated by Dr Sarah Bailey for the CanTest collaborative.(196) These codes are mapped onto ICD-10 and collated into 20 common cancer sites: bladder, breast, cervix, head and neck, kidney, leukaemia, lymphoma, myeloma, oesophagus, pancreas, stomach, testis, uterus, brain, colorectal, lung, ovary, oral, melanoma, prostate and 'other'.

2.3.5. **Autoimmune code-lists**

A full list of autoimmune conditions was developed using the Mackay and Rose Textbook 'The Autoimmune Diseases' (fifth edition).(197) This was cross checked against the American Autoimmune Related Diseases Association list of autoimmune diseases.(198) The list was developed with the aim of optimising sensitivity and inclusivity, with a total of 829 autoimmune-related medcodes identified. However, as there was a lack of clear evidence from the literature about which autoimmune conditions were associated with a raised inflammatory marker,(197) I used a data-driven approach to determine which codes to include in the final list of autoimmune disease. I used the full code-list of 829 medcodes subdivided into 22 autoimmune conditions to develop a binary variable for each of these 22 autoimmune conditions. I then used Chi^2 testing to look for associations between having a code for any of these autoimmune conditions in the two years before or two years after testing and having a raised inflammatory marker at the index date. No associations between raised inflammatory marker and autoimmune skin conditions, psoriasis, coeliac disease, autoimmune neurological disorders, or autoimmune thyroid disease were found. This was also felt to be clinically plausible, as clinicians would be unlikely to use inflammatory marker testing in these conditions. A shortened code-list containing only autoimmune conditions with some evidence of an association with raised inflammatory markers was therefore produced, with 679 medcodes, categorised into 17 subtypes: rheumatoid arthritis, seronegative arthritis, lupus, vasculitis, polymyalgia rheumatica, dermatomyositis, scleroderma, sarcoid, inflammatory bowel disease, cardiovascular autoimmune diseases, type 1 diabetes, haematological autoimmune diseases, Addison's disease, liver autoimmune diseases, pemphigus & pemphigoid, renal autoimmune diseases and respiratory autoimmune diseases.

2.3.6. Infection code-lists

A total of 2,671 medcodes associated with infection were identified and categorised into 16 sites: bacterial upper respiratory tract infections, viral upper respiratory tract infections, lower respiratory tract infections, lower urinary tract infections, pyelonephritis, gastrointestinal infections, diverticulitis, liver infections, biliary infections, genital tract infections, puerperal infections, cardiovascular infections, neurological infections, bone infections, skin infections and 'other'. Subcategories were retained to allow sub-analyses of the types of infection diagnosed subsequent to inflammatory marker testing. I aimed to make this list inclusive, in order to maximise sensitivity, as both viral and bacterial infections were felt to be potentially clinically relevant in the context of a raised inflammatory marker. Definitions of each subtype of infection are shown in **Appendix A**.

Chronic infections were defined as a separate subcategory (overlapping with several infection sites above); this included tuberculosis, HIV, Lyme disease and syphilis. This distinction was made on the basis that these could cause chronic inflammation, and therefore the cut-off for excluding those with pre-existing chronic infections needed to be longer (see section **2.4.5**).

2.4. Data manipulation

2.4.1. Raw Data received from CPRD

Raw data files were received from the CPRD and imported into Stata 15 for analysis. These included the following files:

- Patient file; containing basic demographic and patient registration details.
- Practice file; containing details of each practice.

- Clinical and referral file; these contain all events in the patients' medical history including symptoms, referrals, measurements, and the date that these were recorded, stored as medcodes.
- Consultation file; containing information on the type of consultation.
- Test file; containing all blood tests, reference ranges and dates of testing.
- Therapy file; containing all prescriptions on the GP system.
- Matching file; containing the details of which cases match which controls.
- Case file; containing the index date defined as the date of the first inflammatory marker test in 2014.

The complete dataset contained 199,928 patients: 160,000 patients with inflammatory marker testing in 2014 and 39,928 matched controls. This is because, of the 40,000 randomly selected cases for matching, 72 were unmatched (no suitable age, gender and practice matched patient available in CPRD).

2.4.2. Linked data provided by CPRD

Linked data provided included Cancer Registry Data, ONS Death Data, Basic Hospital Episode Statistics (HES) and patient level Index of Multiple Deprivation (IMD). Of the 199,928 patients identified in CPRD; 88,250 were eligible for linkage to all of these datasets. I did not solely study participants with data linkage in case this introduced bias; for example, limiting to patients eligible for HES linkage would exclude those who had not attended hospital and could therefore introduce bias towards higher rates of disease in the cohort.

Cancer Registry Data

Of the 199,928 patients identified in the CPRD, 110,245 were eligible for linkage to the cancer registry data; of these, 15,399 were identified in the cancer registry

data. Patients in whom it was not possible to obtain linked data were resident outside England, lacked a valid NHS identifier, were registered at a GP practice which had not consented to linkage or were individuals who had personally dissented from linkage. One file was obtained from the cancer registry. This contained the 'epatid'; the encrypted identifier given to patients in CPRD GOLD. It also contained the date of diagnosis and the site of cancer, defined using ICD-10.

ONS Death Registration Data

This data comes from death certification, which is a legal requirement in England and Wales, and therefore provides the most complete source for mortality statistics. ONS Death Registration contains death information for patients registered at a subset of English practices which participate in the CPRD linkage scheme; this linkage was available for 109,966 patients in our cohort. One file was obtained from ONS Death Registry containing the epatid (the encrypted patient identifier) and the date of death.

Index of Multiple Deprivation

Linked data for patient level index of multiple deprivation was available for 110,181 patients. The Index of Multiple Deprivation (IMD) is a composite score, based on patients' postcode, categorised into five equally sized quintiles, where 1 corresponds to the least deprived, and 5 the most deprived. Not all patients in CPRD are eligible for linkage. To be eligible for inclusion patients must be registered in a practice which has consented to take part in the CPRD-linkage scheme (which is currently restricted to practices in England). They also must have a full postcode recorded in their electronic records, and have no record indicating dissent from transmission of personal confidential data to NHS Digital.

Basic Hospital Episode Statistics

HES linked data was available for 93,923 patients. HES linkage is only available in England and only for patients who are admitted to hospital, attend accident and emergency departments or outpatient clinic appointments. Hospital episode statistics (HES) were used to obtain data on ethnicity of the sample, for descriptive purposes. Ethnicity data was therefore only available for the minority of patients who had been to hospital. Ethnicity is coded as white, black Caribbean, black African, black other, Indian, Pakistani, Bangladeshi, other Asian, Chinese, Mixed, other or unknown.

2.4.3. Variables generated – participant characteristics

Raw variables used included 'epatid' which is the unique identifier given to each patient in CPRD. Several new variables were generated to allow analysis. The matching file was used to generate a categorical 'case-control' variable defined as follows:

1- Matched controls: 39,928 controls without inflammatory marker testing in 2014, matched by age, sex and practice to a random sample of cases.
2- Matched cases: 39,928 cases with inflammatory marker testing in 2014 with matched controls.
3- Unmatched cases: 72 cases, of the 40,000 randomly selected for matching, who could not be matched to a control.
4- Cohort only: 120,000 patients with inflammatory marker tests in 2014, without a matched control.

A matching variable was generated to enable case-control pairs to be identified. All date variables were converted into 'elapsed dates' (calculated as the number of days from January 1st, 1960) which allows Stata to calculate time periods by adding and subtracting dates.

CPRD provides year of birth, but not month or day of birth. In order to generate an age variable as a continuous variable day and month of birth were set to 1st

July for all patients. This allowed date of birth to be converted into 'elapsed dates' (calculated as the number of days from January 1, 1960) then subtracted from the index date, in order to impute age at testing. Age was then categorised into 10-year bands (<30, 30-39, 40-49, 50-59, 60-69, 70-79, >80).

2.4.4. Variables generated - index tests

The test file was used to identify all inflammatory marker tests in the sample, based on the following entity types: CRP=280, ESR=273, PV=374. The first inflammatory marker test in 2014 was defined as the index test. Controls who did not have any inflammatory marker testing done were allocated the same index date as their matched case. Inflammatory markers were defined as raised if the test result ('data2' in CPRD) was above the mean upper limit of normal from laboratories within our CPRD sample ('data6' in CPRD). For CRP the mean upper limit of normal was 6.8mg/l, for ease of clinical interpretation this was rounded to 7mg/l; for PV this was 1.72mPa.s. For ESR this was stratified by gender and age (see **Appendix B**).

I chose to use a set threshold for defining a raised inflammatory marker, to increase the clinical applicability of the findings. To explore the impact of differences in laboratory assays I retained a second variable defining a raised inflammatory marker based on an individual laboratories upper limit of normal to allow sensitivity analyses (see **2.5.6**). As well as retaining these categorical variables, the continuous test results for each index test, CRP, ESR and PV ('data6' in CPRD) were retained.

For those who had the same inflammatory marker test coded more than once on the same day (n=231), the highest value was retained. In most cases the duplicate tests were identical suggesting duplicate coding of one test.

For those who had more than one type of inflammatory marker test done on the index date a binary variable for 'any raised inflammatory marker' was produced,

which was positive if *either* CRP, PV or ESR was raised (even if the others were normal).

For patients with subsequent repeat inflammatory markers, I generated a categorical variable to measure the trend in CRP, ESR and PV tests ('CRP trend', 'ESR trend' and 'PV trend'). This measured the trend in CRP, ESR and PV in the first repeat test in the 3 months after the index date. For simplicity I only examined the first repeat test, I did not look at patterns of change over 3rd or 4th repeat tests. Three months was chosen as a pragmatic cut-off after discussions with my clinical supervisors; we felt that tests done more than three months after initial testing were unlikely to be related to the same clinical episode, yet a shorter time-period could miss relevant tests due to potential delay in organising follow up inflammatory marker tests. This categorical variable ('CRP/ESR/PV trend') was defined as follows:

1 – index test in the normal range, and no repeat test done
2 – index test normal and repeat test normal
3 – index test normal and repeat raised
4 – index test raised, no repeat test
5 – index test raised, repeat test normal
6 – index test raised, repeat test stable or lower than index, but still above the normal range
7 – index test raised, repeat test rising

To simplify the analysis for those having several tests (CRP, ESR, PV) at two time-intervals I created a simplified variable 'Repeat IM trend'. This was categorised as follows:

0- Normal inflammatory markers at index date
1- One or more index test raised, no repeat test done
2- One or more index test raised, all repeat inflammatory markers normal
3- One or more index test raised, one or more repeat inflammatory markers raised

2.4.5. **Variables generated – target condition**

The target condition was defined as any new cancer, infection or autoimmune conditions (combined and individually), diagnosed following inflammatory marker testing.

Cancer

All cancer diagnoses, other than non-melanomatous skin cancer, were identified by searching the CPRD clinical and referral files for any of 2,120 cancer-related medcodes (see **2.3.4**). Ascertainment of cancer diagnosis was improved by additionally using linked Cancer Registry Data. A variable 'date difference' was generated to record the time elapsed between the index date and the first dated record of cancer in either CPRD or Cancer Registry Data.

A binary variable for pre-existing cancer, 'cancer_old' was created. Participants were defined as having pre-existing cancer if they had any record of cancer in CPRD or cancer registry in the two years prior to the index date. Historic cancers which had not been coded in the medical record for more than 2 years were felt to be unlikely to be relevant to the index inflammatory marker test, so these were not included in the pre-existing cancer category. Patients with undated diagnosis of cancer were included as having pre-existing cancer, as these undated disease codes tend to represent significant past history of disease.

A binary variable for newly diagnosed cancer was created, which was positive for participants without pre-existing cancer, who had a record of cancer in CPRD or cancer registry in the 1 year after the index date. A second binary variable for new cancers diagnosed in the 2 years after the index date was also generated. For those with more than one diagnosis of cancer in the two years after the index date, only the first was retained. A categorical variable for cancer site was generated and defined as 1-21 for the following cancer sites: bladder, breast, cervix, head and neck, kidney, leukaemia, lymphoma, myeloma, oesophagus, pancreas, stomach, testis, uterus, brain, colorectal, lung, ovary, oral, melanoma, prostate and 'other'.

Autoimmune

Autoimmune diagnoses were identified by searching the clinical and referral files of the CPRD records for any of the 679 autoimmune related medcodes.

A binary variable for pre-existing autoimmune conditions was defined as positive if any of these pre-existing autoimmune conditions were coded in the two-year period before the index date. Patients with undated diagnosis of autoimmune conditions were also defined as having pre-existing autoimmune disease, as these undated disease codes tend to represent significant past history of disease.

A binary variable for 'new autoimmune conditions' was generated for those without pre-existing autoimmune disease, who had one of these autoimmune conditions coded in the 1 year after the index date. A second variable for new autoimmune conditions diagnosed in 2 years was also produced. A categorical variable 'new autoimmune site' was created and defined as 1-17 for the following autoimmune conditions; rheumatoid arthritis, seronegative arthritis, lupus, vasculitis, polymyalgia rheumatica, dermatomyositis, scleroderma, sarcoid, inflammatory bowel disease, cardiovascular autoimmune conditions, type I diabetes, haematological autoimmune diseases, Addison's, liver autoimmune diseases, pemphigus & pemphigoid, renal autoimmune diseases and respiratory autoimmune diseases.

A variable 'date difference' was generated to record the time elapsed between the index date of testing, and the first record of an autoimmune condition in CPRD. For those with more than one diagnosis of autoimmune conditions in the two years after the index date, only the first was retained.

Infections

Infections were identified by searching the clinical and referral file for any of the 2,671 infection-related medcodes.

A binary variable for pre-existing infections was generated and defined as positive if patients were coded as having an infection in the 30 days prior to the index date, on the basis that this could represent monitoring rather than new diagnosis of infection. A shorter time period for pre-existing infections compared to cancer and autoimmune conditions was chosen on the basis that patients could have multiple infections within a one-year time period. Patients with chronic systemic infections (TB, HIV, Lyme and syphilis) in the 2 years before the index date were also defined as having pre-existing infection. A binary variable indicating a new diagnosis of infection was generated for patients without pre-existing infection who had a new infection diagnosis on or in the 30 days after the index date of testing.

A variable 'date difference' was generated to record the time period between the index date of testing and the first coding of the diagnosis of an infection. For those with more than one diagnosis of infection in the 30 days after the index date, only the first was retained.

Mortality

I used two sources of data on death: CPRD recorded death date and ONS Death Registry Data. I generated a categorical variable for death which was positive if the patient was recorded as dead within 1 year of the index date according to either CPRD or ONS. A second variable for death within 2 years of the index date was also generated. Date of death was defined as the earliest of ONS and CPRD date of death. I generated a continuous variable of death interval which was defined as the difference between the date of testing and the date of death.

Any relevant disease

A single binary variable for 'any relevant disease' was generated which was positive if the participant had either a new diagnosis of cancer or autoimmune disease within 1 year of the index date, or an infection in 30 days of the index date.

2.4.6. Additional variables generated

Patient symptoms

In order to identify the main symptoms associated with inflammatory marker testing I searched the clinical and referral file for all medcodes in the 28 days prior to and including the index date. I retained the top 200 medcodes most frequently recorded in this period in patients with raised inflammatory markers, the top 200 medcodes in patients with normal inflammatory markers and the top 200 medcodes in the 40,000 untested control group. The untested group were included as a comparator, to determine whether symptoms were at a higher frequency in the tested group.

These three lists were merged and manually reviewed. All medcodes which did not represent symptoms were dropped, including administrative codes and non-specific codes such as 'had a chat to patient'. This left a total of 57 medcodes, representing the most commonly coded symptoms before the index date. Several of these were very closely overlapping (e.g., 'Tiredness symptom' and 'Tired all the time'). The International Classification of Primary Care (ICPC) was used to categorise these 57 medcodes into 45 distinct symptoms. ICPC provides lists of symptoms and diagnoses, provides synonyms for each symptom, and most importantly, lists what should be excluded from the definition of each symptom. Frequency counts were used to determine the 20 symptoms most frequently occurring in the 28 days before the index date in the total dataset from this list of 45 symptoms. I then used the methods described in section 2.3 to generate

complete code-lists for each of these symptoms. Existing code-lists developed by Dr Sarah Bailey using the same ICPC definitions were used as a starting point. Finally, these symptom code-lists were applied to the full dataset to create a binary variable to signify the presence or absence of each of the 20 symptoms in the 28 days prior to the index date.

Process outcomes: frequency of appointments, referrals and blood tests after inflammatory marker testing

Variables were generated to count the frequency of blood tests, phlebotomy appointments and referrals following inflammatory marker testing.

The CPRD consultation file was used to generate a variable to record the number of consultations in the 6 months after the index date. The consultation file contains a variable 'constype' which defines the type of consultation. All administrative consultations were dropped, as were consultations with a duration of 0 minutes, which were assumed to be bogus. Telephone consultations, face to face consultations and home visits were retained. A date difference variable was generated to measure the elapsed dates between the consultation and index date; those in the 6-month period after the index date were retained. A new variable 'consultation count' was generated which counted the number of consultations in this 6-month period. The same process was followed using the referral file, to count the number of referrals in the 6 months following testing.

Finally, the test file was used to count the number of blood tests in the 6 months following the index date. Two variables were created here; the first counted the number of phlebotomy appointments by counting the number of new test dates per patient in the 6 months following the index date. The second counted the number of individual blood tests per patient in this 6-month period.

2.5. Statistical methods and analysis

The statistical methods are mapped to each of the objectives listed in section **1.9.2**. Analysis was performed using Stata15.(199)

2.5.1. Objective 1: Epidemiology of inflammatory marker testing

Objective 1 was to describe the baseline characteristics in terms of age, gender, socioeconomic status and ethnicity of patients having inflammatory marker tests in primary care.

Summary statistics were calculated using Stata to describe the epidemiology of inflammatory marker testing in terms of the age, gender, ethnicity and socioeconomic status of the tested population. Logistic regression was used to calculate the odds of a raised inflammatory marker by age group, gender and socioeconomic status. Continuous test results were explored using histograms. Scatterplots were used to explore associations between inflammatory markers where more than one test was performed simultaneously.

2.5.2. Objective 2: Inflammatory markers and overall disease outcomes

Objective 2 was to determine the diagnostic accuracy of CRP, ESR and PV, singly and in combination, for relevant disease (infections, autoimmune conditions or cancer) in primary care, and to compare disease incidence in tested versus untested populations.

This objective was further broken down as follows:

2a) To determine the incidence of disease following inflammatory marker testing in primary care

2b) To determine the association between inflammatory marker level and disease incidence

2c) To determine the comparative diagnostic accuracy of CRP, ESR and PV for relevant disease in primary care

2d) To determine whether measuring two inflammatory markers simultaneously increases diagnostic accuracy

Objective 2a + 2b: Incidence of disease

Participants with pre-existing disease were excluded from analysis of disease outcomes, including 4,489 with cancers, 11,129 autoimmune conditions and 9,148 infections. This is because I was interested in the utility of inflammatory markers for diagnosis, not the utility of tests for monitoring purposes or in those with known inflammatory conditions.

In the remaining 174,500 participants, the overall incidence of relevant diseases, including infections, autoimmune diseases and cancers (combined and individually), was calculated in patients with raised inflammatory markers versus those with normal inflammatory markers, and in tested versus untested controls. The incidence of disease in the test positive/raised inflammatory marker group is equivalent to the positive predictive value (PPV).

To aid interpretation and clinical utility, test results and disease outcomes were presented using a test consequences graphic based on a nominal primary care tested population of 1000 people.(184)

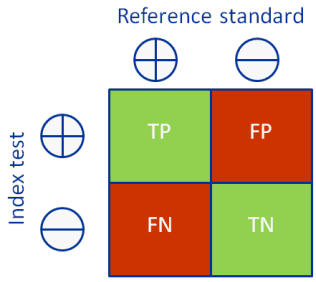
For those with multiple inflammatory marker tests performed simultaneously, disease incidence was calculated in three subgroups: those with multiple raised inflammatory markers, those with discordant results (one raised, one normal), and those with multiple normal results.

For those with a raised inflammatory marker at the index date, disease incidence was calculated in three subgroups: those with no repeat inflammatory marker tests done in the subsequent 3 months, those with normal repeat inflammatory markers, and those with one or more persistently raised inflammatory marker.

Continuous test results for CRP, PV and ESR were categorised into quintiles, and incidence of infections, autoimmune conditions and cancers calculated in each subgroup in order to explore the dose-response relationship between disease incidence and inflammatory marker levels.

Objective 2c: Diagnostic accuracy

For each of the three tests (CRP, ESR and PV), dichotomised test results were cross classified with the target condition 'any relevant disease', allowing sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) to be calculated using the formulae listed in **Figure 3**. The STATA module DIAGT(200) was used to generate 2x2 tables and calculate measures of diagnostic accuracy. Logistic regression was used to calculate diagnostic odds ratios, with and without adjustment for age and gender. As well as looking at 'any relevant disease' I also calculated measures of diagnostic accuracy separately for each of the three disease subtypes: infection, autoimmune conditions and cancer.



True positives	People with relevant disease who have a positive IM test result	TP
True negatives	People without relevant disease who have a negative IM test result	TN
False positives	People without relevant disease who have a positive IM test result	FP
False negatives	People with relevant disease who have a negative IM test result	FN
Sensitivity	Proportion of patients with relevant disease who have a positive IM test result	TP/(TP + FN)
Specificity	Proportion of patients without relevant disease who have a negative IM test result	TN/(FP+ TN)
Positive predictive value (PPV)	Probability that a patient with a positive IM test result has relevant disease	TP/ (TP+FP)
Negative predictive value (NPV)	Probability that a patient with a negative IM test result does not have relevant disease	TN/(FN+TN)

***Figure 3**: 2x2 table showing the cross-classification of index test and reference standard results and overview of measures of accuracy that can be calculated from these data (adapted from Whiting et al (201))*

To address potential concerns that differences in patient mix could lead to biased estimates (for example CRP used preferentially in patients with suspected infection), sensitivity analyses were done on the subgroup of participants who had two tests performed simultaneously to allow head-to-head comparison of diagnostic test accuracy.

Test results were also treated as continuous variables on a log scale (due to their skewed distribution) to assess their predictive value in a logistic regression model, with and without age and gender as additional explanatory variables, calculating the area under curve (AUC), otherwise known as the c-statistic. Sub-analyses compared AUCs for infections, autoimmune conditions and cancers.

Objective 2d: Accuracy of test results in combination

I examined the accuracy of two combinations of inflammatory markers: CRP plus ESR (n=43,820) and CRP plus PV (n=9,575). Only 111 patients had ESR plus PV and 306 had all three tests so I did not examine these test combinations. Where two inflammatory markers were performed simultaneously, I calculated measures of diagnostic accuracy (sensitivity/specificity/PPV/NPV) for two alternative definitions of an overall positive result:

- *Both inflammatory markers raised* (denoted, for example, as CRP + ESR or CRP + PV)
    - Defined as a combined test where both inflammatory markers tested were positive.
- *Either inflammatory marker raised* (denoted, for example, as CRP|ESR or CRP|PV)
    - Defined as a combined test where either of the inflammatory markers tested were positive.

The AUC for test combinations were calculated using a logistic regression model with log transformed test values, including age and gender as co-variates. An interaction term was used in the model due to the associations between inflammatory marker test results. The AUCs for CRP and ESR combined was compared to the better of the two individual tests using the DeLong method(202) generating confidence intervals and p-values. The AUC for CRP and PV combined was likewise compared to the better of the two individual tests.

2.5.3. Objective 3: Symptoms and cascade testing

Objective 3 was to determine the symptomatology of patients with inflammatory marker testing in primary care and measure the consequences of testing in terms of numbers of consultations, blood tests and referrals.

Frequency counts were used to measure frequency of symptoms in the 28 days up to and including the index date in three subgroups: those with normal inflammatory markers, those with raised inflammatory markers and untested controls. This allowed measurement of the symptoms triggering testing, and exploration of which symptoms were more commonly associated with a positive test result.

Frequency counts were used to measure the number of appointments, referrals and blood tests in the 6 months after testing in five subgroups: true positives, false positives, true negatives, false negatives and untested controls.

Comparisons were made between true positives and false negatives and between false positives and true negatives using t test to calculate p-values.

2.5.4. Objective 4: Inflammatory markers and cancer

Objective 4 was to determine the diagnostic accuracy of inflammatory markers for cancer diagnosis in primary care, including stratification by age, gender, inflammatory marker level and cancer type.

This objective was further broken down as follows:

4a) What is the diagnostic accuracy of inflammatory markers, singly and in combination for cancer?

4b) How does this vary by age and gender?

4c) What is the association between inflammatory marker level and cancer?

4d) Which types of cancer are diagnosed following inflammatory marker testing?

The primary analysis reported the one-year cancer incidence (hereafter referred to as 'cancer incidence') for patients with raised versus normal inflammatory markers, and versus untested patients.

I also stratified analysis according to whether multiple inflammatory markers showed concordant or discordant results, whether repeat inflammatory markers were normal or abnormal, and according to symptoms recorded in CPRD in the 28 days prior to testing.

The effect of age and gender on the incidence of disease was examined using a stratified analysis with sub-groups defined by gender and age at index date. The incidence of any relevant disease and corresponding 95% confidence intervals were estimated separately for men and women in each age group, for patients with raised inflammatory markers, normal inflammatory markers and untested controls.

I used logistic regression to examine the dose-response relationship between CRP, ESR and PV test results as continuous variables, and cancer diagnosis as a binary variable, also generating a receiver-operating characteristic (ROC curve). Logistic regression was also used to generate diagnostic odds ratios (DOR), which were adjusted for age and gender.

A fractional polynomial model was used to model the relationship between cancer incidence and inflammatory markers as a continuous predictor, in order to accommodate the non-linearity in the relationship. These analyses were performed separately for men and women.

Frequency count was used to measure the types of cancer diagnosed following inflammatory marker testing, subdivided by gender.

2.5.5. Objective 5: Inflammatory markers and mortality

Objective 5 was to explore the association between inflammatory markers and one-year mortality in primary care.

This objective was further broken down as follows:

> 5a) What is the predictive value of inflammatory markers, singly and in combination for one-year mortality?
>
> 5b) How does this vary by age and gender?
>
> 5c) What is the association between inflammatory marker level and mortality?
>
> 5c) What is the cause of death in patients with a raised inflammatory marker?

The index date was defined as the first date of inflammatory marker testing in 2014, with one-year mortality defined as death within one year of this index date. Date of death was defined as the earlier date of recorded death in either CPRD or ONS death registry. Cause of death was available from death certification data in 3,141 out of 5,512 deaths where ONS linkage was available.

The primary analysis reported the one-year mortality (hereafter referred to as 'mortality'), stratified by age and gender, for patients with raised versus normal inflammatory markers, and versus untested patients. Sub-analyses were stratified according to whether multiple inflammatory markers showed concordant or discordant results, and whether repeat inflammatory markers were normal or abnormal. I used logistic regression to examine the dose-response relationship between CRP, ESR and PV test results as continuous variables, and mortality as a binary variable, also generating a receiver-operating characteristic (ROC curve). Logistic regression was also used to generate diagnostic odds ratios (DOR), with and without adjustment for age and gender.

A fractional polynomial model was used to model the relationship between mortality and inflammatory markers as a continuous predictor, in order to accommodate the non-linearity in the relationship.

Frequency count was used to measure cause of death in those with ONS registry linkage in three groups: untested controls, normal inflammatory markers and raised inflammatory markers.

2.5.6. Sensitivity analyses

Several sensitivity analyses for overall disease outcomes were performed; excluding those with less than one year follow up in CPRD (for example patients moving practice), using the laboratory's own upper limit of normal rather than our own derived thresholds and finally restricting analysis only to those patients eligible for linkages.

2.6. Summary

In this chapter I have outlined the methods of the CPRD study, including the aims and objectives, study design, sources of data, methods of code-list development, stages of data manipulation and finally statistical methods and analyses used. In the next chapter I will present the results of this study, and in the final discussion chapter I will explore the strengths and weaknesses, and implications for research and clinical practice.

CHAPTER 3. QUANTITATIVE RESULTS

3.1. Chapter overview

The overall aim of this chapter is to describe the diagnostic utility of inflammatory markers in primary care for relevant disease, including infection, autoimmune disease, and cancer. The results are mapped onto the quantitative objectives which were listed in **section 1.9.2**. Detailed methodology has been described in **Chapter 2**.

3.2. Objective 1: Epidemiology of inflammatory marker testing

Objective 1 was to describe the baseline characteristics in terms of age, gender, socioeconomic status, and ethnicity of patients having inflammatory marker tests in primary care.

3.2.1. Study sample

From 12,905,266 potentially eligible patients aged >18 in CPRD GOLD in 2014, 3,197,446 (24.8%) have had an inflammatory marker test at some point, of these 463,304 (3.6%) had one or more inflammatory marker test in 2014. Of these 160,000 were selected at random by CPRD. Patients with missing inflammatory marker test results were excluded from the analysis (n=673) as well as 2 patients with test results so abnormal they were considered spurious, leaving 159,375 patients with inflammatory marker tests in this analysis.

3.2.2. **Demographics**

Baseline characteristics of the cohort are shown in **Table 5**. The overall cohort was 61.6% female with a median age of 55.4 years (interquartile range 41.1 - 69.9 years). Compared to the UK adult population, the tested cohort was more likely to be female (62.0% versus 51.3%), of white ethnicity (87.8% versus 85.4%), and from the most affluent socioeconomic quintile (23.1% versus 20.0%).

Table 5: *Characteristics of the tested cohort, compared to the UK population(203) and compared to the population attending GP practices(204)*

Patient Characteristic	**Tested Cohort**	**Untested matched controls**	**UK national population(203)**	**Population attending GP practices(204)**
Gender:	**(159,325)**	**(39,928)**	*****	******
Male	38.0%	38.2%	48.7%	39.8%
Female	62.0%	61.8%	51.3%	60.2%
Age Group:	**(159,325)**	**(39,928)**	*****	******
18-29	10.9%	10.8%	20.2%	12.5%
30-39	11.4%	11.5%	16.4%	13.7%
40-49	15.9%	15.8%	17.8%	14.7%
50-59	17.6%	17.7%	16.3%	14.1%
60-69	17.7%	17.3%	13.9%	16.3%
70-79	15.1%	15.4%	9.4%	15.5%
≥80	11.5%	11.4%	6.0%	13.2%
Postcode IMD Socioeconomic Status*:	**(87,839)**	**(22,063)**		
1 (least deprived)	23.1%	23.9%	20.0%	
2	21.9%	22.2%	20.0%	
3	21.6%	21.4%	20.0%	
4	19.1%	19.0%	20.0%	
5 (most deprived)	14.2%	13.6%	20.0%	
Ethnicity*:	**(75,802)**	**(17,868)**	******	
White	87.8%	87.7%	85.4%	
Non-white	12.3%	12.4%	14.6%	

**IMD linkage and ethnicity data was not available for the complete dataset, hence a lower denominator for these figures shown*

3.2.3. Tests requested

Figure 4 shows the tests requested; 114,198 (71.7%) had a CRP test, 92,325 (58.0%) had an ESR and 15,994 (10.0%) had a PV test. Of those tested, 96,536 (60.6%) had a single inflammatory marker at the index date, and 62,790 (39.4%) had more than one inflammatory marker; mostly CRP and ESR (51,949), followed by CRP and PV (11,107). Four hundred and three had all three inflammatory markers tested simultaneously.

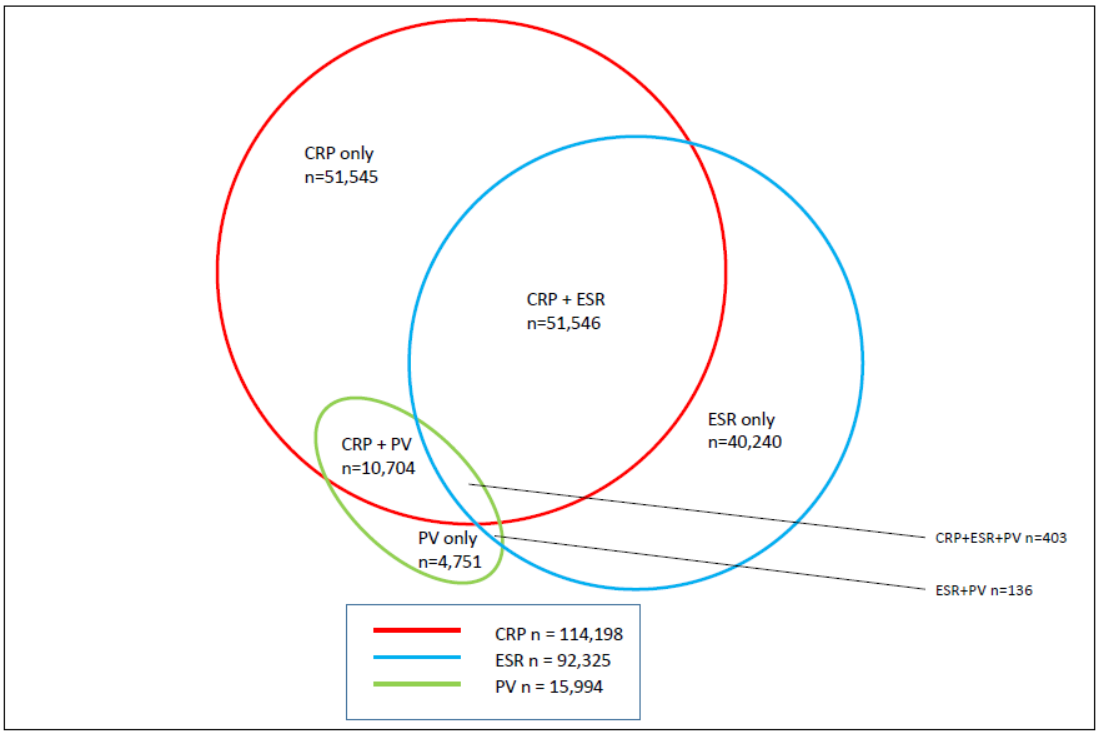


Figure 4: *Inflammatory marker tests requested by GPs in the cohort*

3.2.4. Test results

Of the overall tested cohort, 47,797 (30.0%) had at least one raised inflammatory marker; 25.5% of CRP tests were raised, 25.1% of ESR tests were raised and 28.6% of PV tests were raised. In those with a single inflammatory marker tested, 25.9% had a raised inflammatory marker. In comparison, 36.3% of those with multiple simultaneous tests done had one or more raised inflammatory marker; 14.4% had

concordant raised values and 22.0% had discordant results (one raised, one normal).

Figure 5 shows a scatter plot showing the association between CRP and ESR test results for those who had these two tests performed simultaneously. Whilst an association was seen, the R-squared was only 0.285, meaning that only 28% of the movement in the independent variable can be explained by the dependent variable. **Figure 6** shows a similar association between CRP and PV test results, this time the R-squared was only 0.194. This demonstrates that, whilst there is an association between these inflammatory markers, they are not closely correlated.

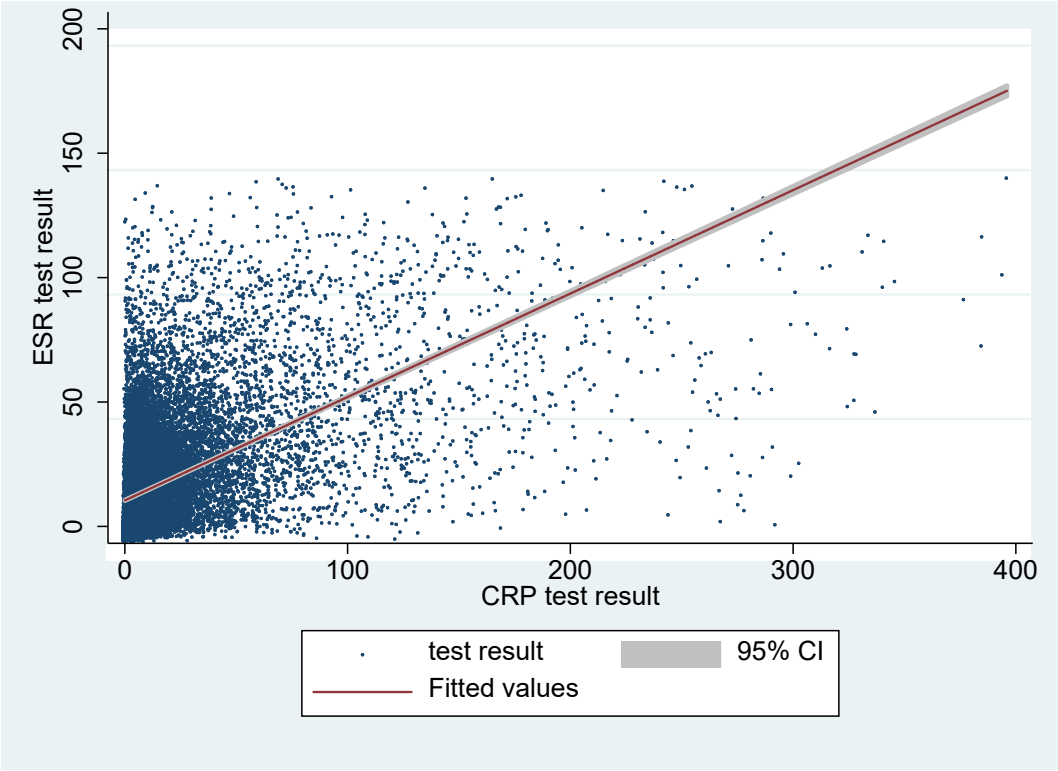


Figure 5: *scatter plot of ESR against CRP test results with fitted linear regression line*

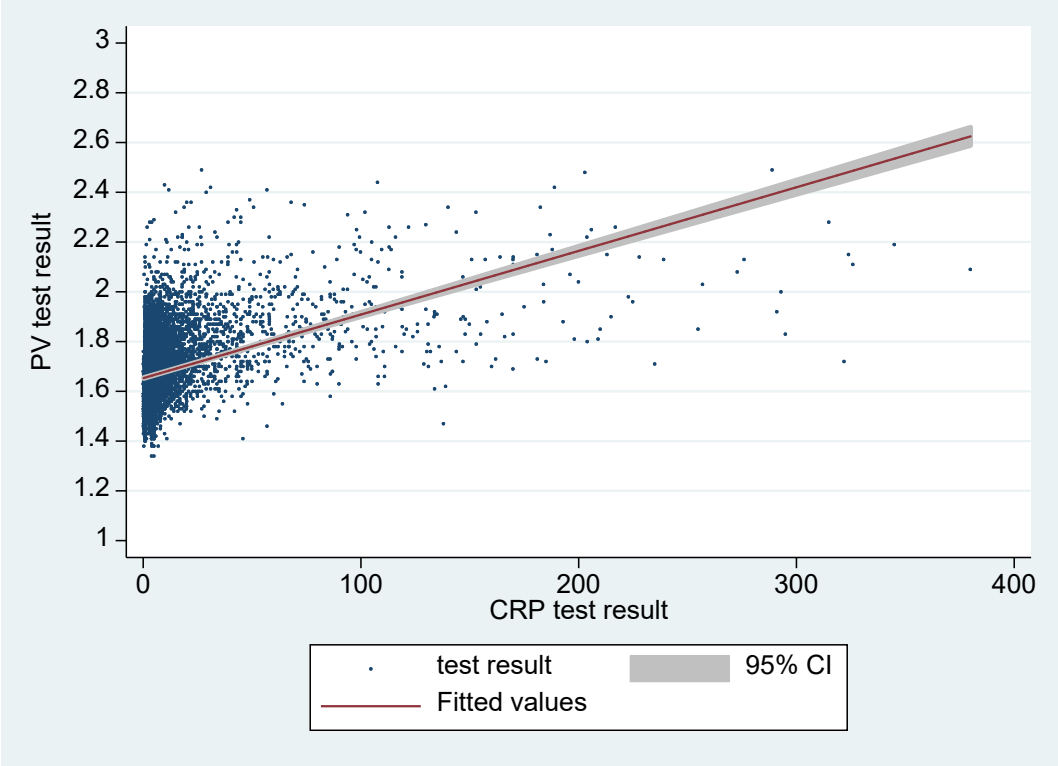


***Figure 6**: Scatterplot of CRP against PV test results with fitted linear regression line*

Mean unadjusted test results by age, gender and index of multiple deprivation are shown in **Table 6**. For all three tests the mean test result increased with increasing age. Mean CRP test results were higher in men, ESR test results were higher in women, for PV no gender difference was seen. Higher mean ESR test results in women are in keeping with the higher upper limit of normal for ESR in women.

Table 7 shows the overall odds of a raised inflammatory marker according to demographic characteristics, with and without adjustment. Raised inflammatory markers were more common amongst the most deprived socioeconomic quintile (adjusted OR 1.39, 95% CI 1.32–1.46, p<0.001), more common amongst women (adjusted OR 1.13, CI 1.10-1.17, p<0.001) and more common with increasing age.

Table 6: *Mean test results according to age, gender, socioeconomic status and ethnicity (unadjusted)*

	Mean test result (SD)		
Patient Characteristic	**CRP**	**ESR**	**PV**
Gender:			
Male	12.4 (30.7)	12.0 (17.0)	1.68 (0.16)
Female	9.9 (27.7)	15.1 (16.2)	1.68 (0.14)
Age Group:			
18-29	7.7 (20.9)	8.3 (10.0)	1.64 (0.11)
30-39	7.7 (20.8)	9.5 (11.3)	1.66 (0.17)
40-49	7.7 (21.6)	10.7 (12.6)	1.66 (0.13)
50-59	8.9 (23.6)	12.7 (14.3)	1.68 (0.13)
60-69	11.6 (29.6)	15.4 (17.6)	1.69 (0.14)
70-79	14.6 (33.4)	18.6 (19.8)	1.70 (0.18)
≥80	18.7 (37.6)	22.9 (22.7)	1.71 (0.16)
IMD Socioeconomic Status:			
1 (least deprived)	10.2 (27.1)	13.2 (16.2)	1.67 (0.12)
2	10.5 (27.4)	13.8 (16.4)	1.68 (0.17)
3	10.6 (27.9)	13.9 (16.9)	1.68 (0.14)
4	11.0 (28.1)	14.0 (16.2)	1.69 (0.13)
5 (most deprived)	11.4 (29.2)	14.1 (16.6)	1.69 (0.14)
Ethnicity*:**			
White	11.6 (29.8)	14.4 (15.8)	1.68 (0.14)
Non-white	8.3 (22.3)	14.4 (17.1)	1.70 (0.16)

Table 7: *Odds of any raised inflammatory marker (CRP, ESR or PV) by age group, gender, ethnicity and socioeconomic status*

	Unadjusted model		**Adjusted***	
	OR	**95% CI**	**OR**	**95% CI**
Gender:				
Male	1.0		1.0	
Female	1.15	1.13 to 1.18	1.13	1.10 to 1.17
Age Group:				
18-30	1.00		1.0	
30-39	1.21	1.15 to 1.27	1.19	1.11 to 1.29
40-49	1.25	1.19 to 1.31	1.24	1.16 to 1.33
50-59	1.47	1.41 to 1.54	1.47	1.38 to 1.58
60-69	1.83	1.76 to 1.92	1.83	1.72 to 1.96
70-79	2.15	2.05 to 2.25	2.13	1.99 to 2.28
≥80	3.00	2.86 to 3.14	2.90	2.71 to 3.11
IMD Socioeconomic Status:				
1 (least deprived)	1.00		1.0	
2	1.10	1.05 to 1.14	1.10	1.05 to 1.15
3	1.15	1.10 to 1.20	1.17	1.11 to 1.22
4	1.23	1.18 to 1.29	1.30	1.24 to 1.37
5 (most deprived)	1.28	1.22 to 1.35	1.39	1.32 to 1.46
Ethnicity:				
White	1.0		1.0	
Non-White	1.00	0.96 to 1.05†	1.12	1.07 to 1.18

**Adjusted for gender, age group, IMD socioeconomic status, and ethnicity*

3.2.5. **Repeat testing**

The overall mean number of repeat inflammatory marker tests performed in the 12 months following the index date was 1.09 (ranging from 0 to 104 repeat tests). This figure counts each additional CRP, ESR or PV test performed simultaneously or sequentially after the index date as one repeat test. For those with normal inflammatory markers at the index date the mean number of repeat tests was 0.80, for those with a single raised inflammatory marker at the index date this increased to 1.32, with discordant inflammatory markers (one raised one normal) this was 1.84 and with multiple raised inflammatory markers there was an average of 2.91 repeat tests. The majority of these repeat tests were done

in the first 3 months after the index date (see **Figure 7**). This shows that raised inflammatory markers are associated with repeat inflammatory marker testing, usually within 3 months of the first test.

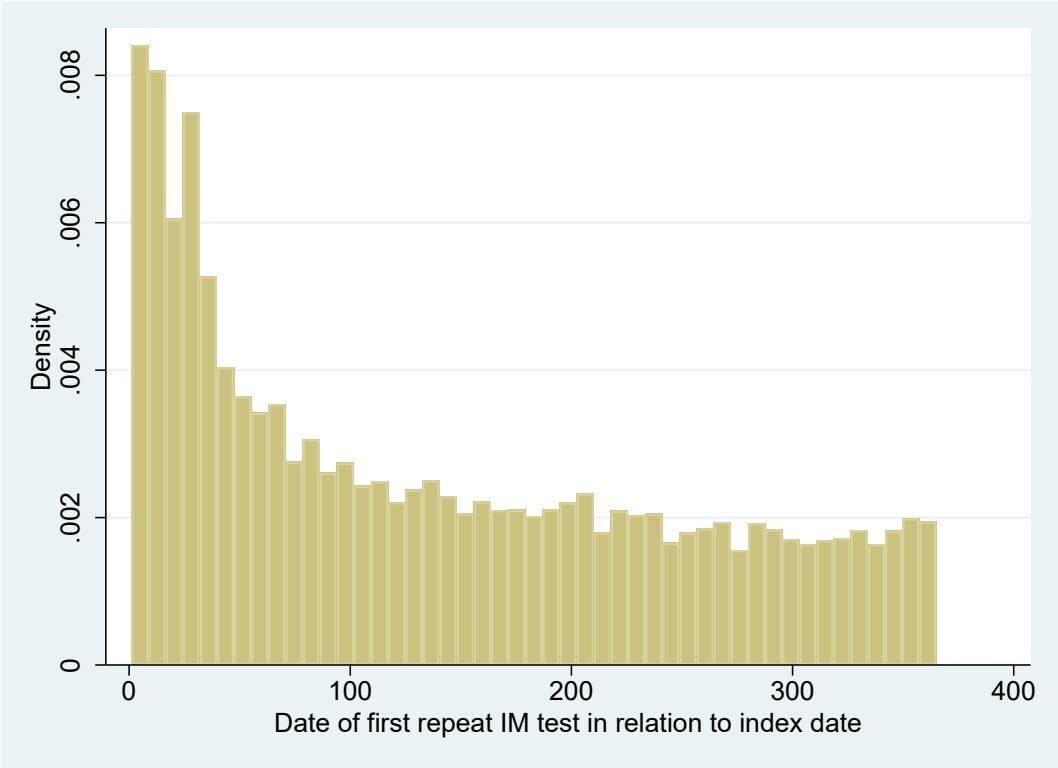


Figure 7: *Histogram showing the number of days between the index and the first repeat inflammatory marker test*

3.3. Objective 2: Inflammatory markers and overall disease outcomes

Objective 2 was to determine the diagnostic accuracy of CRP, ESR and PV, singly and in combination, for relevant disease (infections, autoimmune conditions or cancer) in primary care, and to compare disease incidence in tested versus untested populations.

This objective was further broken down as follows:

> 2a) To determine the incidence of disease following inflammatory marker testing in primary care. This is covered in section 3.3.2 to 3.3.4.
>
> 2b) To determine the association between inflammatory marker level and disease incidence. This is covered in section 3.3.5.
>
> 2c) To determine the comparative diagnostic accuracy of CRP, ESR and PV for relevant disease in primary care. This is covered in section 3.3.6 and 3.3.8
>
> 2d) To determine whether measuring two inflammatory markers simultaneously increases diagnostic accuracy. This is covered in section 3.3.7 and 3.3.9

- Watson J, Salisbury C, Whiting P, Banks J, Pyne Y, & Hamilton W. (2019). Added value and cascade effects of inflammatory marker tests in UK primary care: a cohort study from the Clinical Practice Research Datalink. British Journal of General Practice, 69(684), e470-e478. https://doi.org/10.3399/bjgp19X704321
- Watson J, Jones H, Banks J, Whiting P, Salisbury C, & Hamilton W. (2019). Use of multiple inflammatory marker tests in primary care: using Clinical

Practice Research Datalink to evaluate accuracy. British Journal of General Practice, 69(684), e462-e469. https://doi.org/10.3399/bjgp19X704309

3.3.1. **Study sample**

As the aim of this analysis was to explore new diagnosis of infection, autoimmune disease, or cancer, those with pre-existing disease were excluded (4,489 cancers, 11,129 autoimmune conditions, 9,148 infections) as well as 660 with missing test results and 2 with test results so abnormal they were considered spurious, leaving 136,961 in the final cohort (see **Figure 8**).

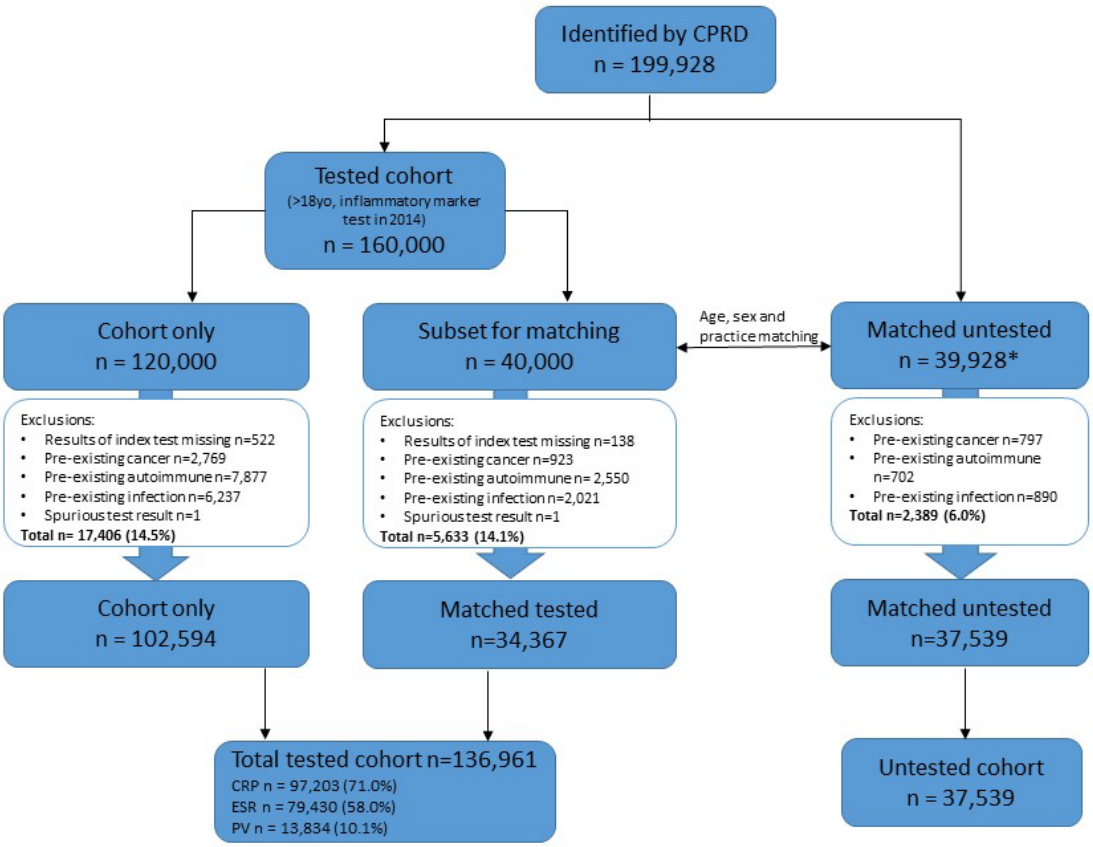


Figure 8: *Flowchart showing exclusions*

**Matched untested consists of 39,928 because of the 40,000 from the cohort who were randomly selected for matching, 78 had no suitable age, sex and practice matched control*

3.3.2. **Overall disease incidence**

The overall disease incidence in the tested cohort was 8.49%: 3.88% infections, 2.75% autoimmune conditions, 2.12% cancers. This compares to 3.44% disease incidence in the untested comparison cohort; 2.02% infections, 0.53% autoimmune conditions and 0.94% cancer. This demonstrates an increased disease incidence in the population selected for inflammatory marker testing. When the tested population were subdivided according to test results, the overall incidence of disease in patients with a raised inflammatory marker (PPV) was 15.0%: 6.3% infections, 5.6% autoimmune conditions, 3.7% cancers (**Table 8**). In comparison, in those with normal inflammatory markers overall disease incidence was 6.0%: 2.9% infections, 1.7% autoimmune disease and 1.5% cancers. The disease incidence in those with normal inflammatory markers is lower than the raised inflammatory marker group, but higher than the untested comparison cohort (**Table 8**).

Types of infection and autoimmune disease diagnosed in those with raised inflammatory markers are shown in **Table 9** and **Table 10**, cancer types are shown in **Table 24**. The most frequent types of infection diagnosed following a raised inflammatory marker were pneumonia, urinary tract infections, upper respiratory tract infections and cellulitis. The most frequent types of autoimmune condition diagnosed following a raised inflammatory marker were polymyalgia rheumatica, rheumatoid arthritis, inflammatory bowel disease and seronegative arthritis.

Table 8: *Disease incidence according to inflammatory marker test results*

	Any relevant disease (%)	**Autoimmune disease (%)**	**Infections (%)**	**Cancer (%)**
Untested (n=37,539)	1,293 (3.44%)	200 (0.53%)	760 (2.02%)	354 (0.94%)
Normal inflammatory markers* (n=98,951)	5,912 (5.97%)	1,652 (1.67%)	2,908 (2.94%)	1,503 (1.52%)
Raised inflammatory markers** (n=38,010)	5,712 (15.0%)	2,121 (5.58%)	2,407 (6.33%)	1,407 (3.70%)

**All inflammatory marker tests normal **one or more inflammatory markers raised*

Table 9: *Types of infection diagnosed following a raised inflammatory marker test*

Type of infection	**Frequency**	**%**
Pneumonia	640	26.6%
Urinary tract infection	529	22.0%
Viral upper respiratory tract infections*	316	13.1%
Cellulitis	212	8.81%
Bacterial upper respiratory tract infections**	181	7.52%
Hepatobiliary infections	72	3.00%
Diverticulitis	54	2.24%
Other gastrointestinal infections	53	2.20%
Pyelonephritis	31	1.29%
Genital infections	28	1.16%
Other***	291	12.1%
Total	2,407	100%

**including laryngitis, pharyngitis, influenza and common cold ** Including mastoiditis, streptococcal pharyngitis, otitis media, tracheitis, tonsillitis, sinusitis ***Infections occurring in <1% merged into 'other' including bone and joint infections, cardiovascular infections, neurological infections and puerperal infections.*

***Table 10**: Types of autoimmune disease diagnosed following a raised inflammatory marker test*

Type of autoimmune disease	Frequency	%
Polymyalgia rheumatica	862	33.5%
Rheumatoid arthritis	716	27.8%
Inflammatory bowel disease	298	11.6%
Seronegative arthritis	204	7.93%
Respiratory autoimmune diseases*	78	3.03%
Haematological autoimmune diseases*	64	2.49%
Renal autoimmune diseases*	46	1.79%
Sarcoidosis	43	1.67%
Scleroderma	43	1.67%
Liver autoimmune diseases	29	1.13
Lupus	26	1.01%
Other**	165	6.4%
Total	**2,574**	**100%**

**For definitions see Appendix A: Autoimmune diseases and infections included in analysis*
***autoimmune conditions comprising <1% merged into 'other'*

Figure 9 and Figure 10 show the test consequences in terms of test results and disease outcomes, simplified to a nominal population of 1000 tested patients, using natural frequencies to aid interpretation and clinical utility. **Figure 9** uses the graphic developed by Whiting *et al*,(184) **Figure 10** presents the same data using a modified Cates plot. Of those with a positive test result, 85.0% had no evidence of infection, autoimmune condition or cancer ('false positives').

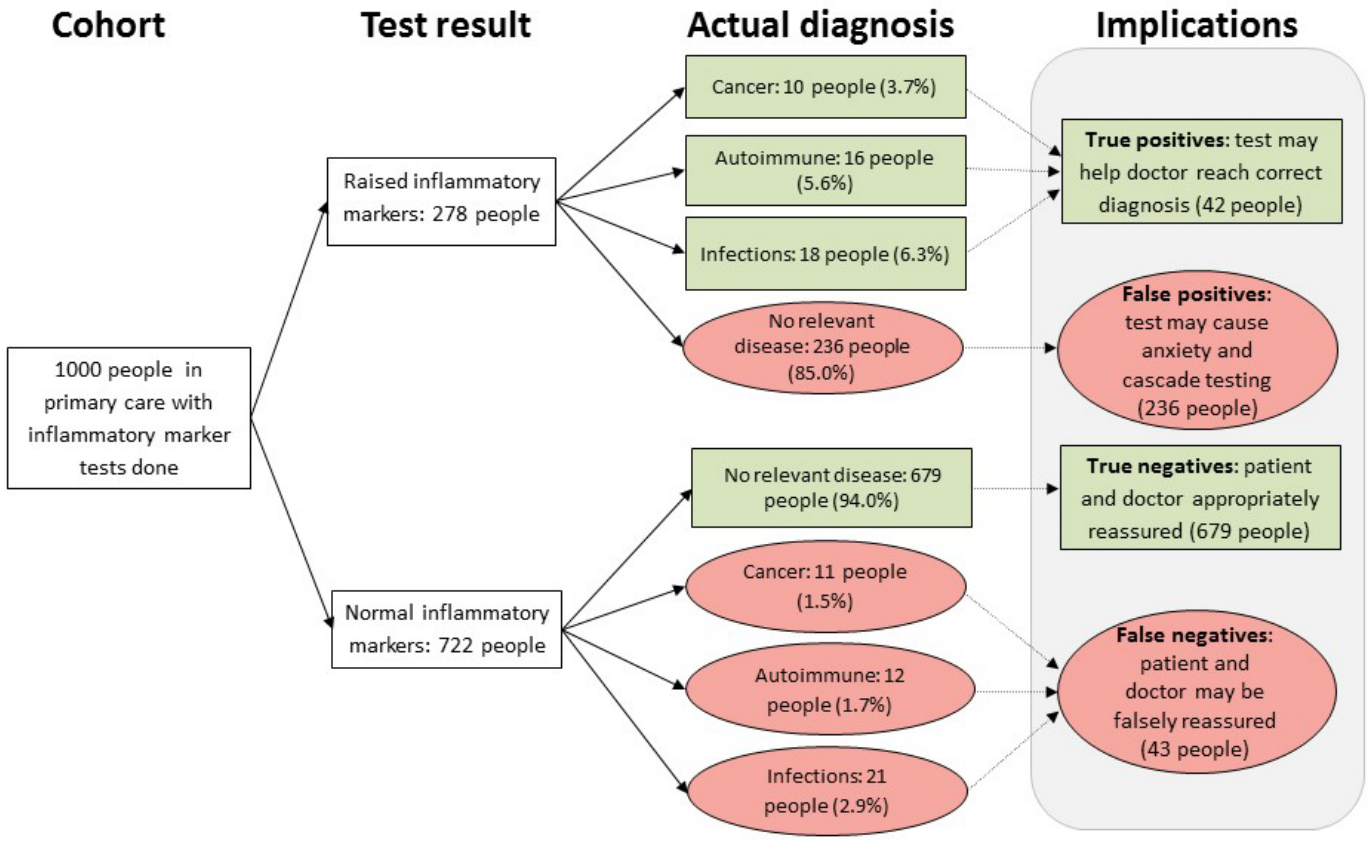


***Figure 9**: Test consequences graphic based on a nominal primary care population of 1000 tested patients*

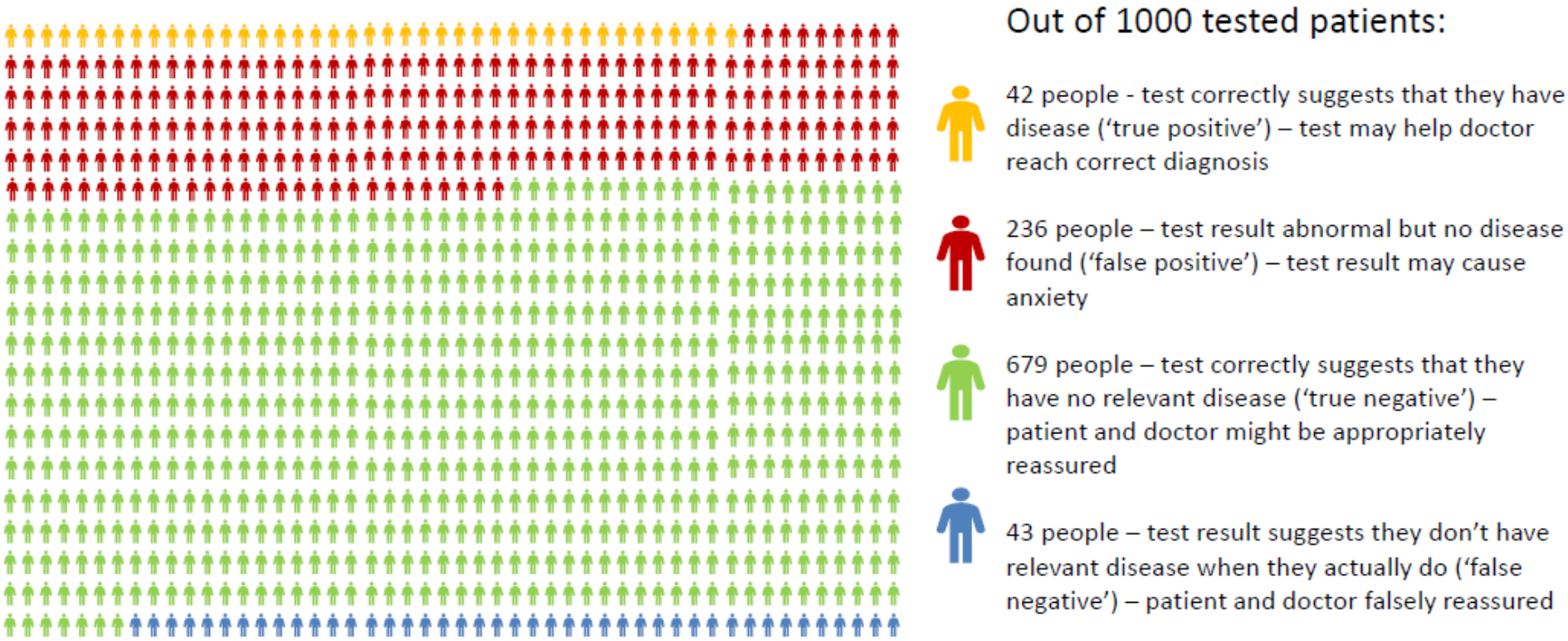


Figure 10: *Test consequences graphic using a modified cates plot*

3.3.3. Disease incidence and multiple simultaneous inflammatory marker tests

Table 11 shows the incidence of disease for those with a single inflammatory marker on the index date, in comparison to those with multiple simultaneous inflammatory markers on the index date. Overall disease incidence was 8.17% in the single tested group compared to 8.98% in the double tested group. In the double tested group, disease incidence was higher in the group with concordant raised values (22.6%), compared to those with a single raised value (10.4%) (**Table 11**). The negative predictive value for any single normal inflammatory marker was 94.0%, compared to 94.1% with multiple normal tests.

Table 11: *Overall disease incidence for single versus double tested patients*

	Test result	Any relevant disease (%)	Autoimmune disease (%)	Infection (%)	Cancer (%)
Single inflammatory marker (n=83,761)	Raised (n=19,932, 23.8%)	2,999 (15.1%)	835 (4.19%)	1481 (7.43%)	804 (4.03%)
	Normal (n=63,829, 76.2%)	3,847 (6.03%)	945 (1.48%)	1,950 (3.06%)	1,043 (1.63%)
Multiple inflammatory markers* (n=53,200)	Multiple raised inflammatory markers (n=6,803, 12.8%)	1,539 (22.6%)	793 (11.7%)	471 (6.92%)	345 (5.07%)
	Discordant results (n=11,275, 21.2%)	1,174 (10.4%)	493 (4.37%)	455 (4.04%)	258 (2.29%)
	All normal (n=35,122, 66.0%)	2,065 (5.88%)	707 (2.01%)	958 (2.73%)	460 (1.31%)

**This includes patients with two and three inflammatory marker tests. If two or more tests were raised patients were included in the 'multiple raised inflammatory markers'; those with only one raised inflammatory marker were included in the 'discordant results' category.*

3.3.4. Disease incidence and repeat testing

Of those with one or more raised inflammatory marker (n=38,010), incidence of disease was highest amongst those with persistently raised inflammatory markers on repeat testing, lower in those with normal inflammatory markers on repeat testing, and lowest amongst those without repeat testing in the next 90 days (**Table 12**). Cancer was the notable exception; those with a normal repeat test had a lower disease risk (2.11%) than those without repeat testing (3.57%).

Table 12: *Disease incidence according to repeat inflammatory marker test results in those with one or more raised inflammatory marker at the index date (n=38,010)*

	Any relevant disease (%)	Infections (%)	Autoimmune (%)	Cancer (%)
No repeat inflammatory markers done (n= 27,874)	3,471 (12.5%)	1,643 (5.89%)	952 (3.42%)	994 (3.57%)
Repeat inflammatory markers normal (n=3,314)	615 (18.6%)	249 (7.51%)	321 (9.69%)	70 (2.11%)
One or more raised inflammatory marker on repeat testing (n=6,822)	1,626 (23.8%)	515 (7.55%)	848 (12.4%)	343 (5.03%)

3.3.5. Disease incidence and inflammatory marker levels

Incidence of disease increased with rising inflammatory marker levels in a dose-response relationship (**Figure 11**). With CRP ≥100mg/L (n=1552), 501 (32.3%) developed one or more relevant diseases: 109 (7.02%) cancers, 99 (6.38%) autoimmune conditions and 317 (20.2%) infections. With ESR ≥100mm/h (n=389), 141 (36.3%) developed one or more relevant diseases: 59 (15.2%) cancers, 60 (15.4%) autoimmune conditions and 36 (9.25%) infections. With PV ≥2mPas

(n=276), 81 (29.4%) developed one or more relevant diseases: 30 (10.9%) developed cancer, 38 (13.8%) developed autoimmune conditions and 15 (5.43%) developed infections.

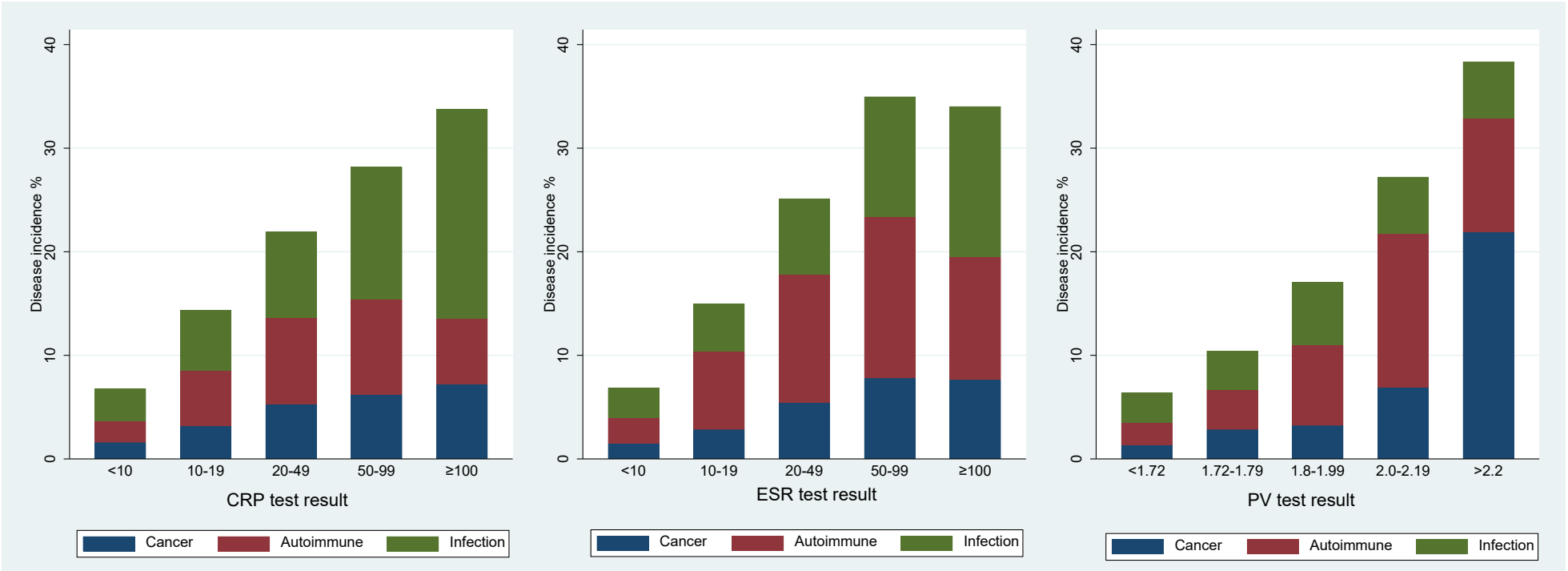


Figure 11: *Incidence of relevant disease in relation to test result for CRP, ESR and PV*

For the small number (<0.5%) with more than one disease outcome; cancer superseded autoimmune disease, which superseded infections

3.3.6. Diagnostic accuracy of single inflammatory marker tests: sensitivity, specificity, PPV, NPV

Table 13 summarises the detailed performance characteristics of each of the three tests (CRP, ESR, PV) for the primary outcome of any relevant disease as well as for each of the three disease outcomes separately. All three tests had sensitivities below 50% for any relevant disease. CRP, ESR and PV were broadly similar in terms of sensitivity, specificity, PPV and NPV.

In order to address the issue of confounding by indication, sensitivity analyses were done to calculate the measures of diagnostic accuracy limited to subgroups with both CRP and ESR done together (**Table 14**) and with CRP and PV together (**Table 15**); minimal differences to overall results were found. Very few patients had both ESR and PV done together so it was not possible to directly compare these tests.

Table 13: *Performance characteristics of inflammatory markers for relevant disease*

	True positives (n)	False positives (n)	True negative (n)	False negative (n)	Sensitivity	Specificity	PPV	NPV	DOR** (unadjusted)	DOR (adjusted for age & sex)
CRP n=97,203										
Any relevant disease	3,947	18,745	69,797	4,714	45.6% (44.5-46.6)	78.8% (78.6-79.1)	17.4% (16.9-17.9)	93.7% (93.5-93.9)	3.12* (2.98-3.26)	2.86* (2.73-2.99)
Infection	1,780	20,912	72,279	2,232	44.4% (42.8-45.9)	77.6% (77.3-77.8)	7.84% (7.5-8.2)	97.0% (96.9-97.1)	2.76* (2.59-2.94)	2.74* (2.57-2.92)
Autoimmune conditions	1,404	21,288	73,063	1,448	49.2% (47.3-51.1)	77.4% (77.2-77.7)	6.19 (5.88-6.51)	98.1% (97.9-98.2)	3.33* (3.09-3.59)	3.05* (2.83-3.29)
Cancer	915	21,777	73,360	1,151	44.3% (42.1-46.5)	77.1% (76.8-77.4)	4.03 (3.78-4.30)	98.5% (98.4-98.5)	2.68* (2.45-2.92)	2.04* (1.86-2.23)
ESR n=79,430										
Any relevant disease	2,780	15,589	57,221	3,840	42.0% (40.8-43.2)	78.6% (78.3-78.9)	15.1% (14.6-15.7)	93.7% (93.5-93.9)	2.66* (2.52-2.80)	2.43* (2.30-2.56)
Infection	918	17,451	59,251	1,810	33.7% (31.9-35.5)	77.3% (77.0-77.5)	5.0% (4.69-5.32)	97.0% (96.9-97.2)	1.72* (1.59-1.87)	1.67* (1.54-1.82)
Autoimmune conditions	1,285	17,084	59,872	1,189	51.9% (50.0-53.9)	77.8% (77.5-78.1)	7.0% (6.63-7.37)	98.0% (97.9-98.2)	3.79* (3.49-4.10)	3.40* (3.14-3.69)
Cancer	696	17,673	60,120	941	42.5% (40.1-45.0)	77.3% (77.0-77.6)	3.79% (3.52-4.08)	98.5% (98.4-98.6)	2.52* (2.28-2.78)	2.07* (1.87-2.30)
Plasma viscosity n = 13,989										
Any relevant disease	536	3,242	9,439	617	46.5% (43.6-49.4)	74.4% (73.7-75.2)	13.9% (12.9-15.1)	93.9% (93.4-94.3)	2.53* (2.23-2.86)	2.32* (2.05-2.62)
Infection	183	3,595	9,767	289	38.8% (34.3-43.3)	73.1% (72.3-73.9)	4.84% (4.18-5.58)	97.1% (96.8-97.4)	1.72* (1.42-2.08)	1.68* (1.39-2.04)
Autoimmune conditions	234	3,544	9,834	222	51.3% (46.6-56.0)	73.5% (72.8-74.3)	6.19% (5.45-7.01)	97.8% (97.5-98.1)	2.92* (2.42-3.53)	2.65* (2.19-3.21)
Cancer	141	3,637	9,922	134	51.3% (45.2-57.3)	73.2% (72.4-73.9)	3.73% (3.15-4.39)	98.7% (98.4-98.9)	2.87* (2.206-3.65)	2.38* (1.86-3.03)

**p<0.001 **DOR=Diagnostic odds ratio: the ratio of the odds of test being positive in disease relative to the odds of positivity in the non-disease*

Table 14: *Comparison of measures of diagnostic accuracy in the subgroup with CRP and ESR tests performed simultaneously (n=43,820)*

Test	True positive (n)	False positive (n)	True negative (n)	False negative (n)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	DOR unadjusted (95% CI)	DOR adjusted* (95% CI)
CRP	1,769	7,725	32,118	2,208	44.5% (42.9-46.0)	80.6% (80.2-81.0)	18.6% (17.9-19.4)	93.6% (93.3-93.8)	3.33 (3.11-3.56)	3.05 (2.85-3.26)
ESR	1,733	8,815	31,028	2,244	43.6% (42.0-45.1)	77.9% (77.5-78.3)	16.4% (15.7-17.2)	93.3% (93.0-93.5)	2.72 (2.54-2.91)	2.49 (2.32-2.66)

**DOR=diagnostic odds ratio, adjusted for age and gender*

Table 15*: Comparison of measures of diagnostic accuracy in the subgroup with CRP and PV tests performed simultaneously (n=9,575)*

Test	True positive	False positive	True negative	False negative	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	DOR unadjusted (95% CI)	DOR adjusted* (95% CI)
CRP	369	1,862	6,896	448	45.2% (41.7-48.7)	78.7% (77.9-79.6)	16.5% (15.0-18.2)	93.9% (93.3-94.4)	3.05 (2.63-3.53)	2.84 (2.45-3.30)
PV	393	2,222	6,536	424	48.1% (44.6-51.6)	74.6% (73.7-75.5)	15.0% (13.7-16.5)	93.9% (93.3-94.5)	2.73 (2.36-3.15)	2.53 (2.18-2.93)

**DOR=diagnostic odds ratio, adjusted for age and gender*

3.3.7. Diagnostic accuracy of test results in combination: sensitivity, specificity, PPV and NPV

For two simultaneous inflammatory marker tests, by definition, sensitivity and specificity vary depending on how the results are interpreted. Where tests were performed simultaneously, measures of diagnostic accuracy (sensitivity/specificity/PPV/NPV) were calculated for two alternative definitions of an overall positive result:

Both inflammatory markers raised (denoted, for example, as CRP + ESR or CRP + PV) Defined as a combined test where both inflammatory markers tested were positive.
Either inflammatory marker raised (denoted, for example, as CRP|ESR or CRP|PV) Defined as a combined test where either of the inflammatory markers tested were positive.

Results are shown in **Table 16**. If an overall positive result was defined as *both* inflammatory markers raised (e.g. CRP + ESR), then PPVs were higher, specificity was increased, but at the price of lower sensitivity, compared to using any single test.

If the combined test was defined as *either* inflammatory marker raised (e.g. CRP|ESR), then sensitivity increased but specificity fell compared to any single test. This led to fewer false negatives or reduced risk of missed diagnoses but a markedly increased frequency of false positives; for example, CRP alone generated false positives in 19.3% of those tested, compared to 32.5% false

positives for CRP|PV. The maximum sensitivity was 60.6% for the test combination CRP|PV. The maximum specificity was 89.3% for the test combination CRP + ESR.

Table 16: *Measures of diagnostic accuracy of two inflammatory marker tests in combination for relevant disease (infection, autoimmune disease or cancer)*

Test	Number of patients	True positive	False positive	True negative	False negative	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	DOR unadjusted (95% CI)	DOR adjusted*** (95% CI)
CRP + ESR*	43,820	1,277 (2.91%)	4,269 (9.74%)	35,574 (81.2%)	2,700 (6.16%)	32.1 (30.7-33.6)	89.3 (89.0-89.6)	23.0 (21.9-24.2)	93.0 (92.7-93.2)	3.94 (3.66-4.24)	3.56 (3.31-3.84)
CRP | ESR**	43,820	2,225 (5.08%)	12,271 (28.0%)	27,572 (62.9%)	1,752 (4.00%)	56.0 (54.4-57.5)	69.2 (68.8-69.7)	15.4 (14.8-16.0)	94.0 (93.8-94.3)	2.85 (2.67-3.05)	2.61 (2.44-2.79)
CRP + PV	9,575	267 (2.79%)	974 (10.2%)	7,784 (81.1%)	550 (5.74%)	32.7 (29.5-36.0)	88.9 (88.2-89.5)	21.5 (19.3-23.9)	93.4 (92.9-93.9)	3.88 (3.30-4.56)	3.56 (3.02-4.19)
CRP | PV	9,575	495 (5.17%)	3,110 (32.5%)	5,648 (59.0%)	322 (3.36%)	60.6 (57.1-64.0)	64.5 (63.5-65.5)	13.7 (12.6-14.9)	94.6 (94.0-95.2)	2.79 (2.41-3.23)	2.59 (2.24-3.01)

**CRP + ESR; positive test defined as both CRP and ESR positive. **CRP | ESR; positive test defined as either CRP or ESR positive ***DOR=diagnostic odds ratio, adjusted for age and gender*

3.3.8. Comparative accuracy of inflammatory marker tests: area under curve (AUC)

Pairwise comparisons of AUC for CRP versus ESR (**Table 17; Figure 12**) and CRP versus PV (**Table 18; Figure 13**) were done in those with multiple simultaneous tests, in a model containing age and gender as additional explanatory variables. Receiver operating curves comparing CRP and ESR are shown in **Figure 12**, curves comparing CRP and PV are shown in **Figure 13**. Very few patients had ESR and PV simultaneously, so it was not possible to directly compare these two tests.

Comparing CRP with ESR, CRP had a slightly higher AUC for infection (**Table 17**: AUC 0.617 versus 0.589, p=<0.0001). There was no significant difference in the AUC of CRP and ESR for diagnosis of cancer or autoimmune disease (including the main subtypes of autoimmune disease: rheumatoid arthritis, seronegative arthritis, polymyalgia rheumatica or inflammatory bowel disease).

Comparing CRP and PV, a higher AUC of CRP for infection was found (**Table 18**: AUC 0.638 versus 0.597, p=0.004), with no difference for cancer (p=0.49) or autoimmune disease (p=0.97). The accuracy of CRP versus PV was not compared for sub-types of autoimmune disease due to the smaller sample size for PV tests.

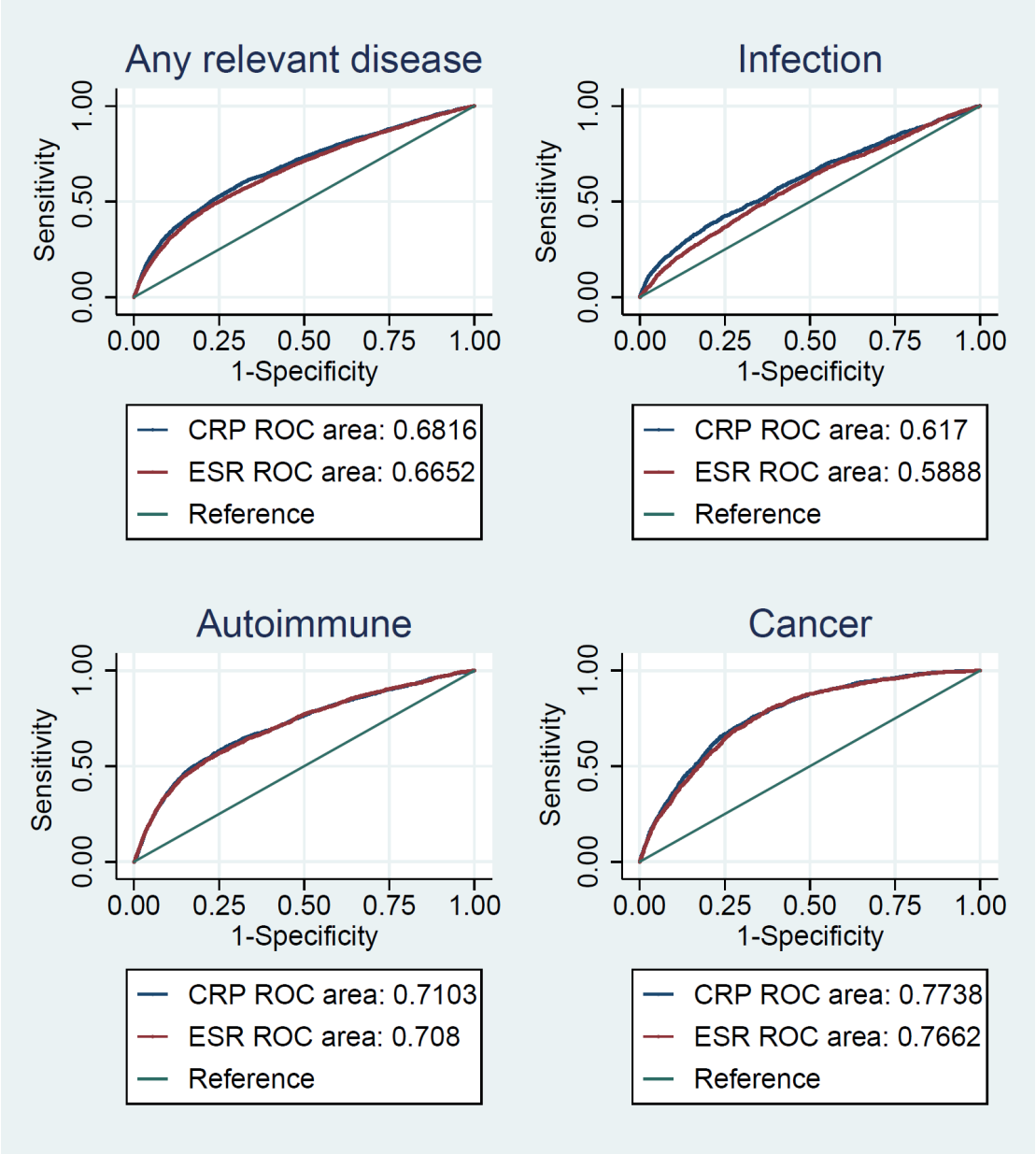


Figure 12: *ROC curves comparing CRP (blue lines) and ESR (red lines) in patients with both CRP and ESR performed simultaneously (n=43,820) for any relevant disease, infection, autoimmune disease and cancer*

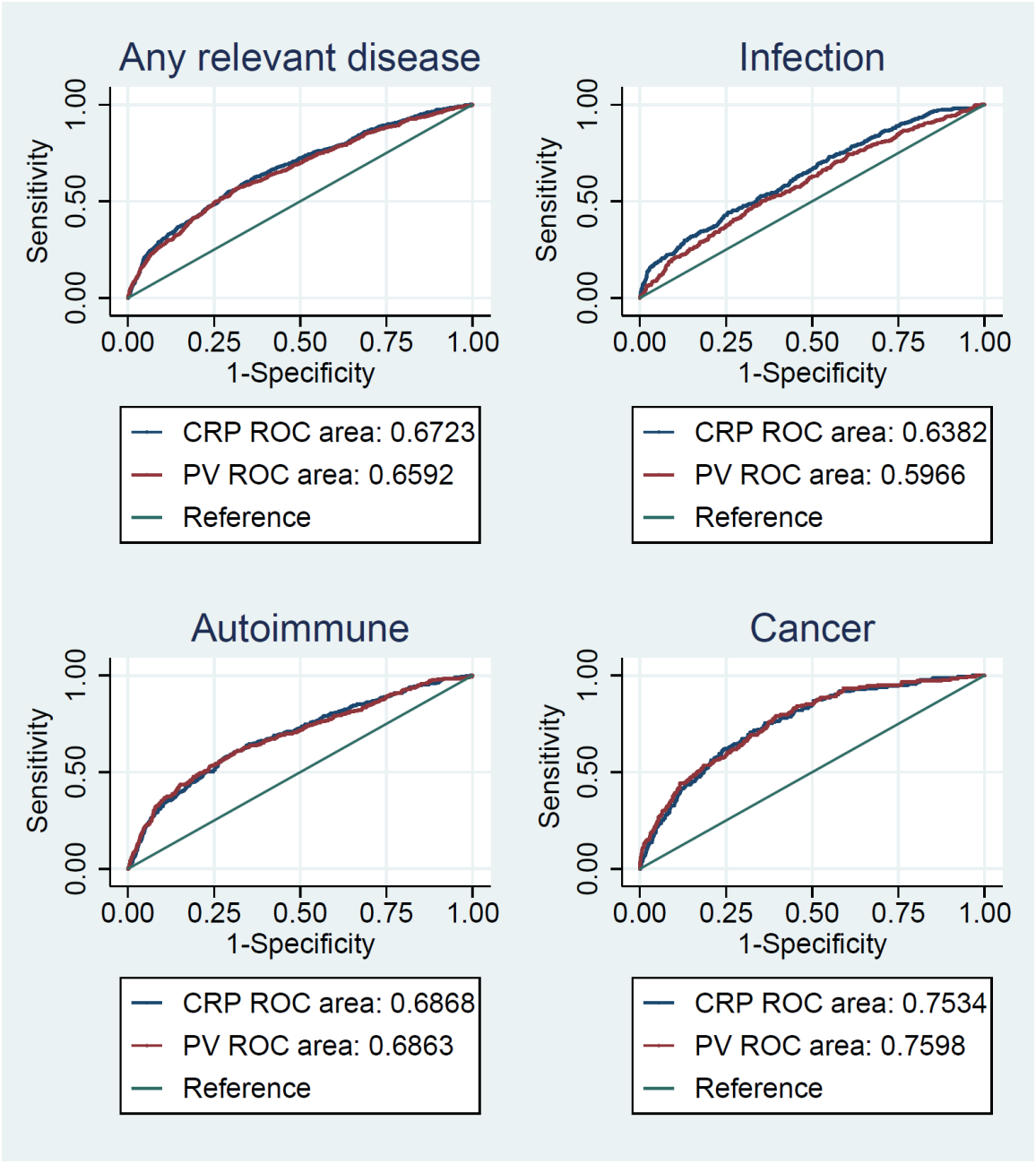


Figure 13: *ROC curves comparing CRP (blue lines) and PV (red lines) in patients with both tests performed simultaneously (n=9,575) for any relevant disease, infection, autoimmune disease and cancer*

3.3.9. Comparative accuracy of test results in combination: area under curve (AUC)

The accuracy of CRP and ESR in combination, was compared to the better of the two individual tests (**Table 17**). There was no improvement in AUC for CRP and ESR in combination for infection. The combined test CRP plus ESR gave an increase of 0.014 in the AUC for autoimmune disease ($p<0.001$) and 0.003 increase in AUC for cancers ($p=0.006$). Whilst these p-values reach conventional thresholds for statistical significance, the differences are unlikely to be of a magnitude to be clinically significant. The combined test did not increase the AUC for polymyalgia rheumatica, seronegative arthritis or inflammatory bowel disease, and led to a small increase of 0.009 in the AUC for rheumatoid arthritis ($p=0.007$).

Similarly, the combination of CRP and PV together gave no improvement in AUC, compared to the better of the two individual tests, for infection or cancer (**Table 18**). The combined test CRP and PV gave an increase of 0.022 in the AUC for autoimmune disease, although the p-value was 0.09 the magnitude of the difference seems unlikely to be clinically significant.

Table 17: *Comparison of AUC amongst those with both CRP and ESR performed simultaneously (n=43,820)*

Disease outcome (n)**	**CRP AUC* (95% CI)**	**ESR AUC (95% CI)**	**P value for CRP v ESR**	**CRP and ESR AUC (95% CI)**	**P value for combined versus better single test**
Any relevant disease (3,977)	0.682 (0.672-0.690)	0.665 (0.656-0.674)	<0.001	0.688 (0.678-0.697)	<0.001
Infections (1,565)	0.617 (0.601-0.632)	0.589 (0.574-0.603)	<0.001	0.619 (0.604-0.634)	0.018
Autoimmune conditions (1,663)	0.710 (0.697-0.724)	0.708 (0.695-0.721)	0.68	0.724 (0.710-0.737)	<0.001
Cancer (882)	0.774 (0.759-0.788)	0.766 (0.752-0.781)	0.017	0.777 (0.763-0.791)	0.006
Polymyalgia rheumatica /giant cell arteritis (476)	0.882 (0.88- 0.90)	0.872 (0.86-0.89)	0.099	0.887 (0.874-0.900)	0.11
Rheumatoid arthritis (557)	0.691 (0.670-0.712)	0.690 (0.669-0.711)	0.89	0.700 (0.679-0.721)	0.007
Seronegative arthritis (151)	0.700 (0.653-0.746)	0.686 (0.638-0.734)	0.51	0.706 (0.659-0.753)	0.54
Inflammatory bowel disease (223)	0.698 (0.660-0.737)	0.691 (0.653-0.730)	0.45	0.701 (0.662-0.740)	0.51

**AUC = area under the receiver operating curve, where AUC=0.5 is equivalent to no diagnostic utility and AUC=1 is perfect diagnostic accuracy. AUC was calculated using logistic regression modelling with test result(s) on a log scale and age and gender as additional explanatory variables **Where disease sub-types were examined, diseases other than the specified condition reported were classified as non-diseased*

Table 18: *Comparison of overall test performance amongst those with both CRP and PV performed simultaneously (n=9,575)*

Disease outcome** (n)	CRP AUC* (95% CI)	PV AUC (95% CI)	P value for CRP v PV	CRP and PV AUC (95% CI)	P value for combined versus better single test
Any relevant disease (n=817)	0.672 (0.652-0.692)	0.659 (0.640-0.679)	0.17	0.686 (0.667-0.706)	0.004
Infection (n=325)	0.638 (0.608-0.670)	0.597 (0.564-0.628)	0.004	0.639 (0.608-0.669)	0.28
Autoimmune conditions (n=338)	0.687 (0.657-0.717)	0.686 (0.655-0.717)	0.97	0.709 (0.679-0.739)	0.009
Cancer (n=183)	0.753 (0.720 – 0.787)	0.760 (0.726-0.793)	0.49	0.764 (0.731-0.797)	0.17

**AUC = area under the receiver operating curve, where AUC=0.5 is equivalent to no diagnostic utility and AUC=1 is perfect diagnostic accuracy. AUC was calculated using logistic regression modelling with test result(s) on a log scale and age and gender as additional explanatory variables*

***Where disease sub-types were examined, diseases other than the specified condition reported were classified as non-disease*

3.3.10. **Sensitivity analyses**

In order to test the robustness of the data, several additional sensitivity analyses were performed. Firstly, analysis was restricted to those with ≥1 year of follow up, to ensure that loss to follow up was not introducing significant bias. Secondly, analysis was repeated using the laboratory specified upper limit of normal to define a raised inflammatory marker (rather than the mean upper limit of normal see **2.4.4** methods chapter).

These sensitivity analyses gave minor differences in results which were not clinically meaningful (**Table 19**).

***Table 19**: Overall incidence of relevant disease in those with raised inflammatory markers, normal inflammatory markers and untested controls, in selected subgroups*

Subgroup	**Any relevant disease % (95% CI)**		
	Raised inflammatory markers	**Normal inflammatory markers**	**Untested**
Full cohort (n=174,500)	15.0% (14.7-15.4)	5.97% (5.83-6.12)	3.44% (3.26-3.63)
Restricting to patients with ≥1 year follow up in CPRD (n=141,319)	15.8% (15.4-16.2)	6.22% (6.05-6.39)	3.56% (3.36-3.77)
Using laboratory specified upper limit of normal (n=174,500)	14.3% (14.0-14.7)	5.95% (5.80-6.10)	3.44% (3.26-3.63)

3.4. Objective 3: Symptoms and cascade testing

Objective 3 was to determine the symptomatology of patients with inflammatory marker testing in primary care and measure the consequences of testing in terms of numbers of consultations, blood tests and referrals.

3.4.1. Symptoms associated with inflammatory marker testing

Table 20 shows the most common symptoms in the 28 days before testing, ordered according to whether they were relatively more common in patients with normal inflammatory markers (top) or more common in those with raised inflammatory markers (bottom). Broadly the symptoms could be categorised into non-specific symptoms, abdominal symptoms, joint symptoms and infective symptoms. Non-specific symptoms such as tiredness, dizziness and low mood were relatively more common in the test negative compared to test positive groups, indicating that these non-specific symptoms are less likely to generate raised inflammatory markers. In comparison, infective symptoms such as cough, UTI and chest infection were more likely to be associated with a raised inflammatory marker.

Table 20: *Most frequently recorded symptoms in the 28 days before inflammatory marker testing*

	Untested		Normal inflammatory markers		Raised inflammatory markers	
Symptom	**n**	**%**	**n**	**%**	**n**	**%**
Tiredness	82	0.21	6390	5.73	2148	4.43
Dizziness	126	0.32	2156	1.93	755	1.56
Headache	98	0.25	2498	2.24	965	1.99
Low mood	111	0.28	964	0.86	390	0.8
Low back pain	138	0.35	2079	1.87	857	1.77
Back pain	156	0.4	2202	1.98	964	1.99
Abdominal pain	166	0.42	7132	6.40	3232	6.67
Chest pain	100	0.26	1922	1.72	872	1.8
Rash	145	0.37	1663	1.49	797	1.64
Joint pain	52	0.13	2515	2.26	1215	2.51
Pain generalised	87	0.22	1841	1.65	1011	2.09
Diarrhoea	64	0.16	2297	2.06	1266	2.61
Shoulder pain	123	0.31	1103	0.99	631	1.3
Throat symptoms	102	0.26	1018	0.91	598	1.23
Knee pain	182	0.47	1599	1.43	916	1.89
Nausea & vomiting	56	0.14	1171	1.05	720	1.49
Cough	496	1.27	3361	3.02	2336	4.82
Malaise	27	0.07	1005	0.9	720	1.49
UTI	189	0.48	1291	1.16	1057	2.18
Chest infection	130	0.33	720	0.65	804	1.66

Symptoms are ordered according to whether they were relatively more common in patients with normal inflammatory markers (top) or more common in those with raised inflammatory markers (bottom)

3.4.2. Diagnostic activity after initial inflammatory marker test

Table 21 shows the blood tests, appointments, and referrals in the 6 months after testing for 'true positive', 'false positive', 'true negative' or 'false negative' groups, plus untested controls. Test positives were defined as individuals with any one or more raised inflammatory markers at the index date; disease positive was defined as any relevant disease (infection, autoimmune conditions or cancers). Follow on blood tests, appointments and referrals were higher in the false positives compared to the true negatives. Both groups consist of tested patients without subsequent pathology, the main difference being the inflammatory marker result. Based on this, for 1000 inflammatory marker tests performed, one would expect 236 false positives, associated with an additional 710 GP appointments, 229 phlebotomy appointments and 24 referrals in the 6 months following testing.

Table 21: *Diagnostic activity in the 6-month period after testing*

True positives = people with a positive test who develop relevant disease; false negatives = people with a negative test who develop relevant disease; false positives = people with a positive test with no relevant disease; true negatives = people with a negative test with no relevant disease

	Mean number of appointments per person (95% CI)**	**Mean number of phlebotomy appointments in 6 months**	**Mean number of total tests requested**	**Mean number of referrals per person**
True positives n=5,712	15.0* (14.7-15.3)	4.26* (4.15-4.37)	43.7* (42.3-45.2)	0.86* (0.83-0.89)
False negatives n=5,912	12.0* (11.7-12.2)	3.23* (3.14-3.32)	30.0* (28.8-31.2)	0.78* (0.75-0.81)
False positives n=32,298	10.3* (10.2-10.4)	2.78* (2.75-2.82)	24.3* (23.9-24.7)	0.62* (0.61-0.64)
True negatives n=93,039	7.29* (7.2-7.3)	1.81* (1.80-1.83)	13.7* (13.5-13.8)	0.52* (0.51-0.52
Untested controls n=33,486	4.80 (4.74-4.86)	1.14 (1.12-1.16)	9.66 9.47-9.85)	0.24 (0.24-0.25)

**p<0.001 – comparing true positives to false negatives and comparing false positives to true negatives **Includes face to face consultations, home visits and telephone consultations*

3.5. Objective 4: Inflammatory markers and cancer

Objective 4 was to determine the diagnostic accuracy of inflammatory markers for cancer diagnosis in primary care, including stratification by age, gender, inflammatory marker level and cancer type. This objective was further broken down as follows:

4a) What is the diagnostic accuracy of inflammatory markers, singly and in combination for cancer? This is covered in section **3.5.2** to **3.5.6**

4b) How does this vary by age and gender? This is covered in section **3.5.7**

4c) What is the association between inflammatory marker level and cancer? This is covered in section **3.5.8**

4d) Which types of cancer are diagnosed following inflammatory marker testing? This is covered in section **3.5.9**

The results described below have been published in the following Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license:

- Watson J, Salisbury C, Banks J, Whiting P, & Hamilton W. (2019). Predictive value of inflammatory markers for cancer diagnosis in primary care: a prospective cohort study using electronic health records. British Journal of Cancer, 120(11), 1045-1051. https://doi.org/10.1038/s41416-019-0458-x

3.5.1. Study sample

Out of the 160,000 patients with inflammatory markers tested in 2014 and 39,928 untested matched controls (described in 2.2.2), only those with a pre-existing cancer diagnosis (n=4,489), and patients who had missing or spurious

inflammatory marker test results (n=662) were excluded from this analysis (**Figure 14)**. This was to ensure the results were generalisable to a primary care population without a current cancer diagnosis, but not limited to those without other co-morbidities.

After exclusions, the inflammatory marker cohort contained 155,646 patients; of these 111,440 (71.6%) had a CRP test, 90,478 (58.1%) had an ESR test, and 15,670 (10.1%) had a PV test. Altogether, 61,545 (39.5%) had more than one test performed on the index date, mostly CRP and ESR (50,522), followed by CRP and PV (10,494). Of the tested cohort, 46,092 (29.6%) had at least one raised inflammatory marker.

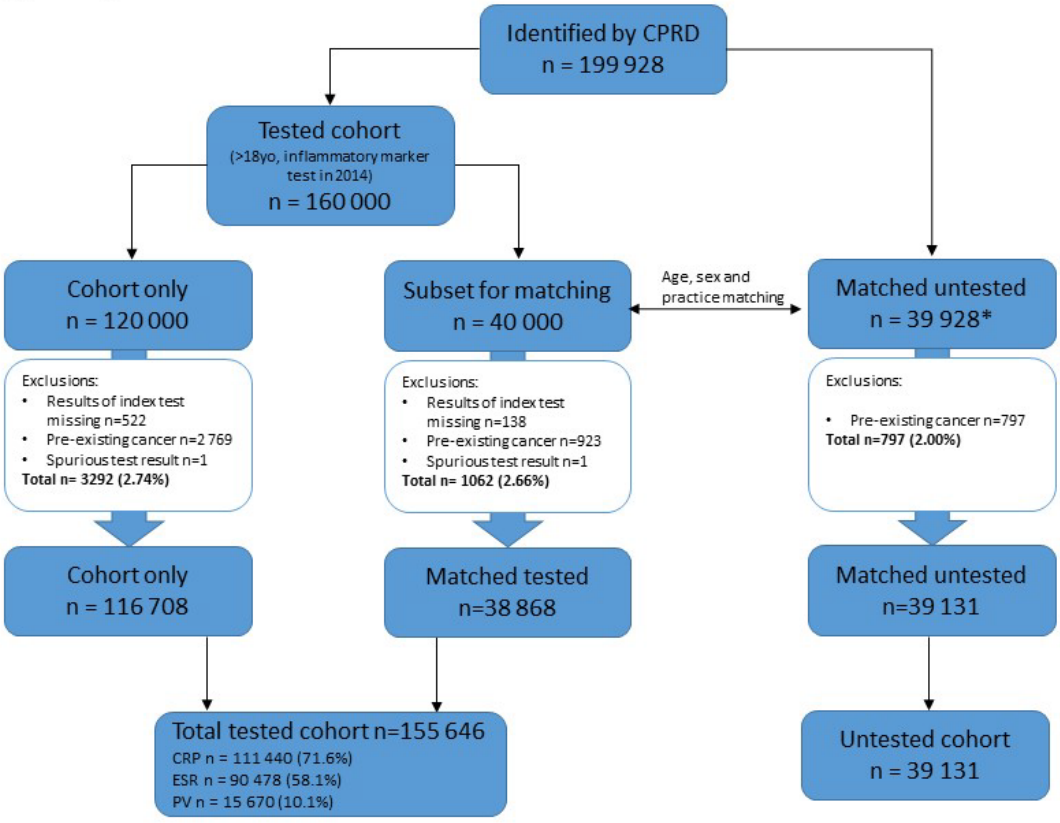


***Figure 14**: Flowchart showing exclusions for cancer analyses*

Linked data from the English Cancer Registry was obtained by the CPRD for 110, 245 patients. Patients in whom it was not possible to obtain linked data were resident outside England, lacked a valid NHS identifier, were registered at a GP

practice which had not consented to linkage or were individuals who had personally dissented from linkage.

3.5.2. Overall cancer incidence according to test results

Table 22 shows the overall cancer incidence according to test results. In patients with a raised inflammatory marker, cancer incidence was 3.53% (95% CI 3.37 to 3.70), compared to 1.50% (1.43 to 1.58) with normal inflammatory markers ($p<0.001$). The untested cohort had a cancer incidence of 0.97% (0.87 to 1.07). Most cancer diagnoses were made within one year of testing (see **Figure 15**), with no evidence of significantly increased cancer risk in the second year after a raised inflammatory marker compared to untested controls.

Table 22: *Cancer incidence according to test result*

	Raised inflammatory markers*	**Normal inflammatory markers****	**Untested**
Number of patients (n)	46 092	109 554	39 131
Cancers diagnosed in 1 year (n)	1 629	1 648	379
1-year cancer incidence % (95% CI)***	3.53 (3.37-3.70)	1.50 (1.43-1.58)	0.97 (0.87-1.07)
Second year cancer incidence % (95% CI)	1.07 (0.97-1.16)	0.77 (0.72-0.82)	0.96 (0.86-1.05)

**One or more inflammatory marker raised **All inflammatory markers tested normal*
****Equivalent to positive predictive values (PPV) for the test positive group*

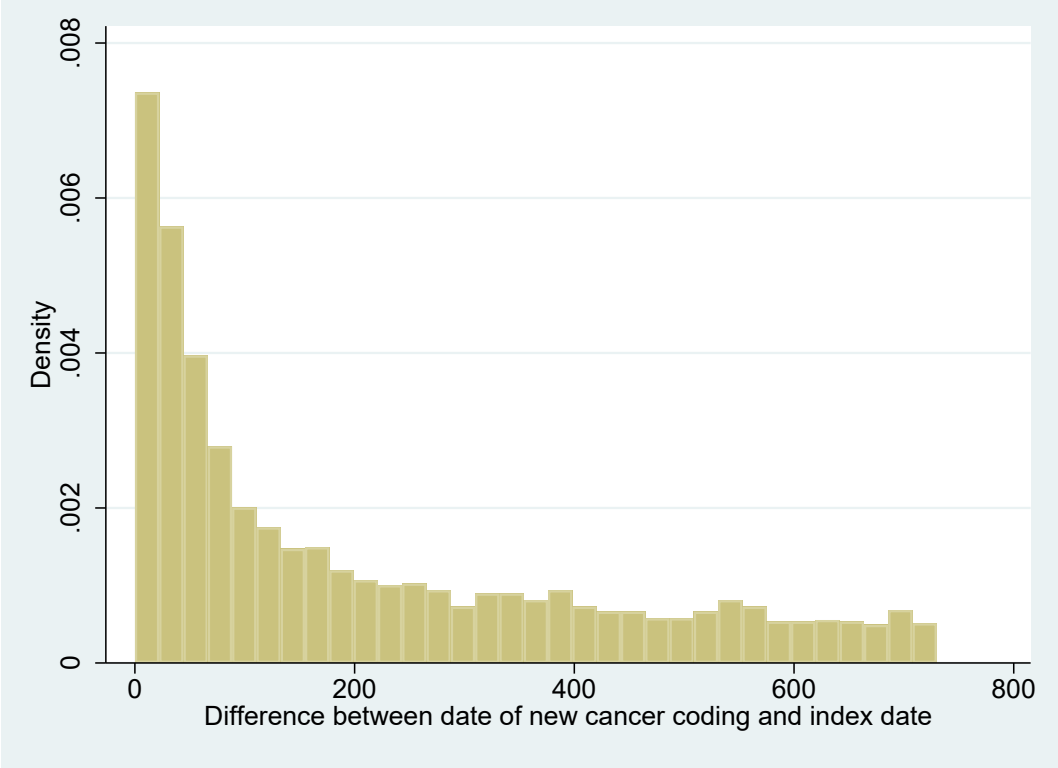


Figure 15: *Histogram showing time interval in days between index date and the date of cancer diagnosis for those with a raised inflammatory marker*

3.5.3. Cancer incidence and multiple inflammatory marker testing

For the 61,545 who had more than one inflammatory marker test done together, if both were normal (n=39 368; 64.0%), the cancer incidence was 1.28% (1.17 to 1.39); with one raised and the other normal (n=13 472; 21.9%) the cancer incidence was 2.27% (2.02 to 2.52); if both were raised (n=8 705; 14.1%) then cancer incidence was 4.71% (4.26 to 5.16).

3.5.4. Cancer incidence and repeat testing

When patients with a raised inflammatory marker (n=46 092) had a second inflammatory marker test taken within 90 days (n=13 873), the cancer incidence

was greatest if the second test result was further increased, at 6.86% (5.73 to 7.99) for CRP, 5.04% (4.01 to 6.08) for ESR and 4.13% (2.00 to 6.26) for PV. If the second test was lower than the first, but still above the normal range, the cancer incidence was 4.35% (3.66 to 5.04) for CRP, 3.55% (2.86 to 4.24) for ESR and 3.28% (1.79 to 4.78) for PV. If the repeat test was normal the cancer incidence fell to 1.98% (1.49 to 2.47) for CRP, 2.49% (1.73 to 3.26) for ESR, and 1.32% (0.34 to 2.29) for PV. **Figure 16** shows this as a flowchart of cancer incidence according to repeat test results for CRP, the most frequently used of the three inflammatory markers.

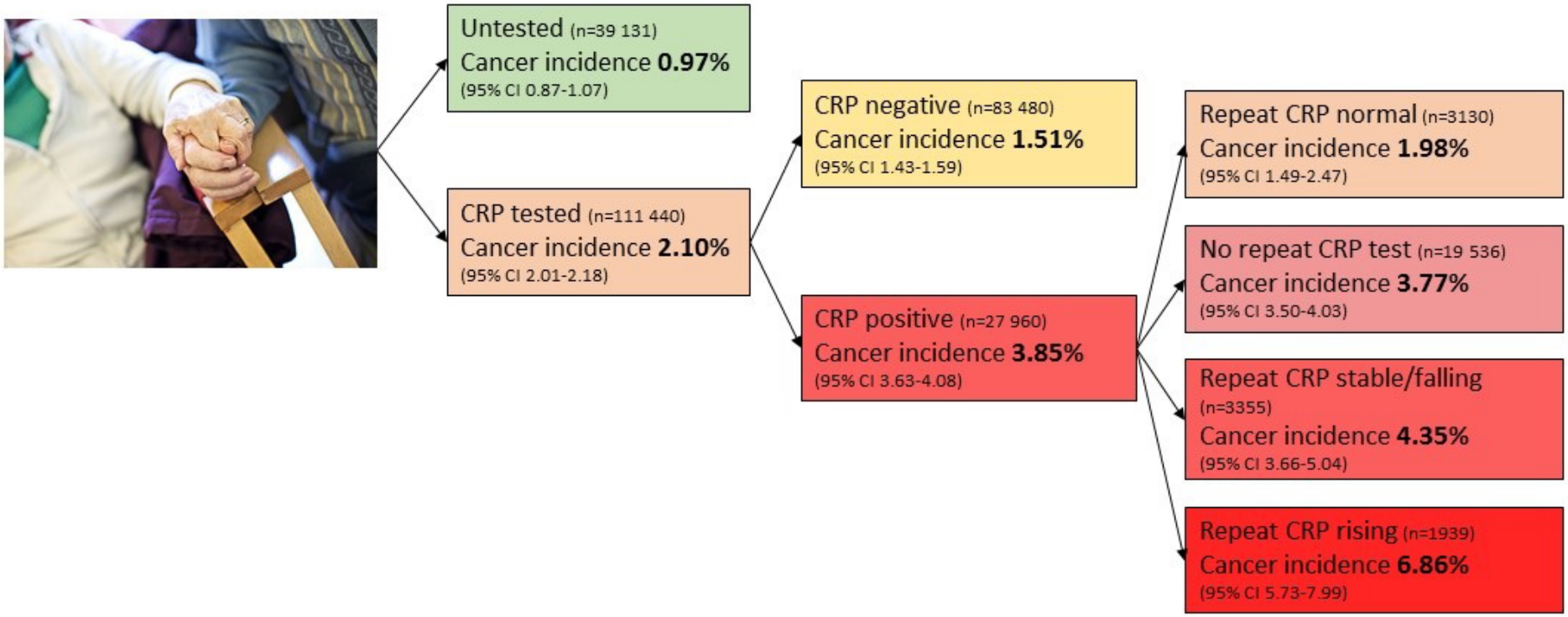


***Figure 16**: Flowchart of one-year cancer incidence according to CRP test results*

3.5.5. Measures of diagnostic accuracy: sensitivity, specificity, PPV and NPV

For each of the three tests, sensitivities, specificities, PPV, NPV and DOR for a positive result are shown in **Table 23**. All three inflammatory markers had a low sensitivity of <50%. Sensitivity and specificity were comparable for the three tests.

Table 23: *Performance characteristics of inflammatory marker tests for cancer*

	True positives (n)	False positives (n)	True negatives (n)	False negatives (n)	Sensitivity	Specificity	PPV	NPV	DOR** (unadjusted)	DOR (adjusted for age + gender)
CRP (n=111,440)	1,077	26,883	82,219	1,261	46.1% (44.0-48.1)	75.4% (75.1-75.6)	3.85% (3.63-4.08)	1.51% (1.43-1.59)	2.29* (2.12-2.46)	1.79* (1.66-1.93)
ESR (n=90,478)	810	21,628	66,993	1,047	43.6% (41.4-45.9)	75.6% (75.3-75.9)	3.61% (3.37-3.85)	1.54% (1.44-1.63)	1.98* (1.83-2.15)	1.75* (1.61-1.90)
PV (n=15,670)	153	4,289	11,073	155	49.7% (44.0-55.4)	72.1% (71.4-72.8)	3.44% (2.91-3.98)	1.10% (1.16-1.60)	1.69* (1.43-1.99)	1.45* (1.22-1.71)

**p<0.0001 **Diagnostic odds ratio*

3.5.6. Measures of diagnostic accuracy: area under curve (AUC)

A logistic regression model containing age and gender had an AUC of 0.736, compared to 0.747 for a full model containing age, gender and CRP test result as a continuous variable ($p<0.001$); 0.759 with age, gender and ESR ($p<0.001$); and 0.760 with age gender and PV ($p<0.001$).

3.5.7. Effects of age and gender

Breakdown by age and gender (**Figure 17**) shows that cancer incidence increases with age and is higher in men. Women over 60 with a raised inflammatory marker and men over 50 with a raised inflammatory marker, have a cancer incidence of 4.22% (3.90 to 4.45) and 6.44% (6.00 to 6.88) respectively, which exceeds the NICE 3% threshold for urgent cancer referral. For women under the age of 60 with a raised inflammatory marker, cancer incidence was 1.20% (1.03 to 1.37); for men under 50 with a raised inflammatory marker it was 1.03% (0.72 to 1.33). Patients with normal inflammatory markers have a cancer incidence which exceeds that of untested controls but is lower than the raised inflammatory marker group. In particular, men over 60 with normal inflammatory markers have a cancer incidence 3.88% (3.60 to 4.16), above the NICE 3% threshold.

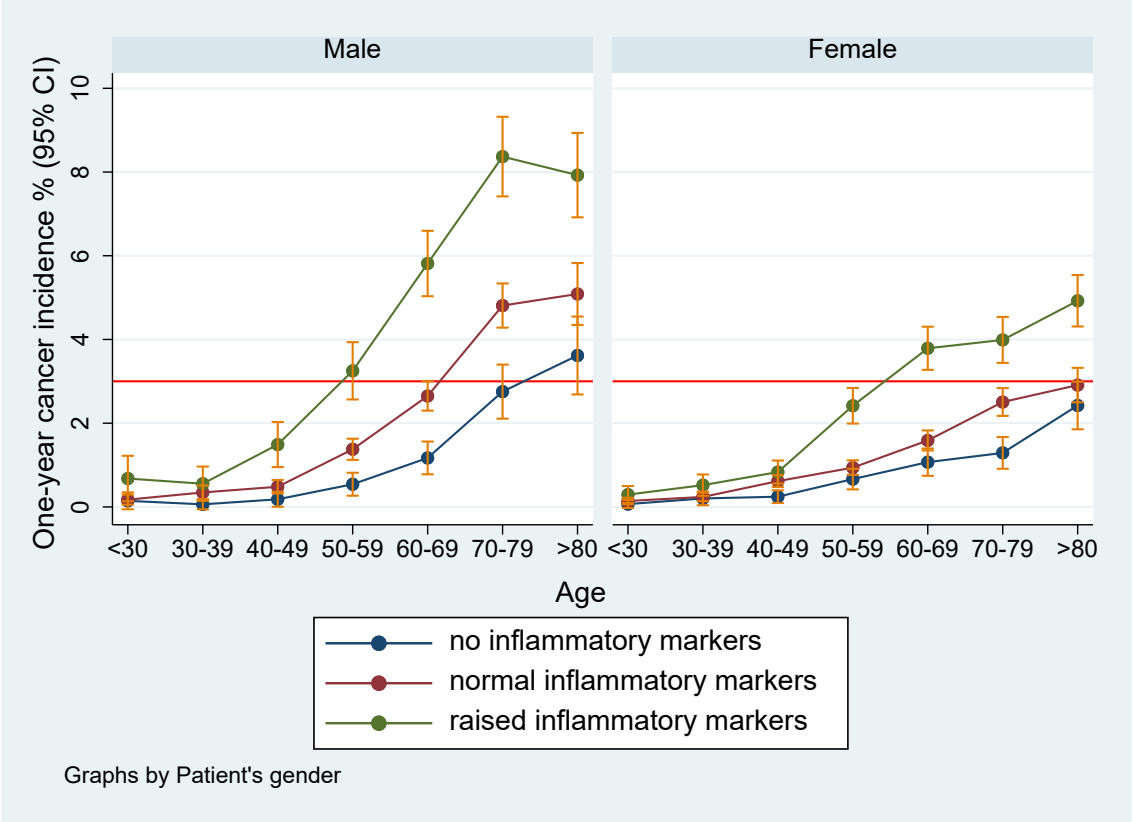


***Figure 17**: One-year cancer incidence, stratified by age, gender and inflammatory marker test result. The red line represents the 3% threshold used by NICE for urgent cancer investigation or referral.*

3.5.8. Cancer incidence and inflammatory marker levels

The incidence of cancer increased with rising inflammatory markers with a dose-response relationship; see **Figure 18**. Out of 506 people with ESR≥100, 69 (13.6%) developed cancer in 1 year; with CRP ≥100 (n=1,983), 135 (6.81%) developed cancer; with PV ≥2.0mm/h (n=342), 31 (9.06%) developed cancer.

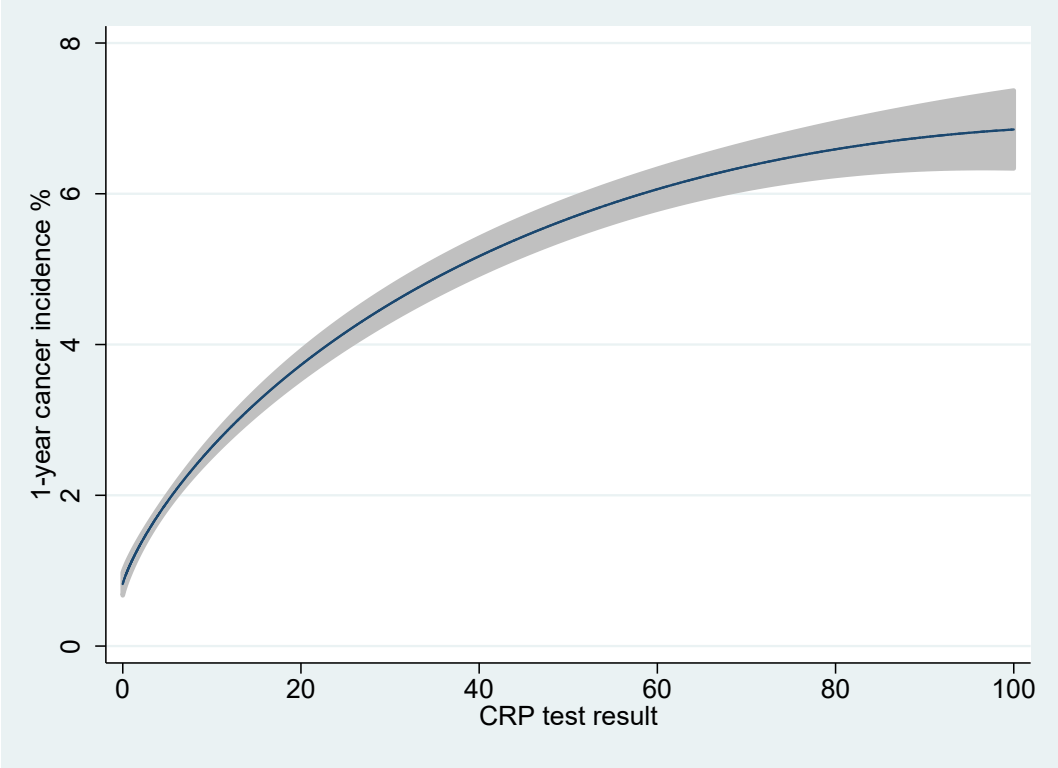


***Figure 18**: Polynomial logistic regression of cancer incidence against CRP test result as a continuous variable*

3.5.9. Cancer sites

The types of cancer are shown in **Table 24**; the cancer sites broadly reflect overall cancer incidence in 2014 from National Cancer Registry figures, apart from breast cancer and prostate cancer, which are notably less frequent in the raised inflammatory marker group. Myeloma contributed only 45 out of 2,145 cancers in the raised inflammatory marker group; sensitivity analysis demonstrated minimal difference in overall results when these were excluded (cancer incidence in the raised inflammatory marker group 3.45% excluding myeloma, versus 3.53% overall; **Table 26**).

***Table 24**: Types of cancer diagnosed by gender in those with raised inflammatory markers compared to types of cancer diagnosed nationally in England in 2014.(205)*

	Male		Female	
	Raised inflammatory markers	**National incidence**	**Raised inflammatory markers**	**National incidence**
Cancer site	**% (n)**	**%**	**% (n)**	**%**
Bladder	3.91 (41)	4.14	1.59 (17)	1.56
Breast	0.19 (2)	0.22	15.78 (169)	31.56
Cervix	-	-	1.12 (12)	1.77
Head and neck	1.34 (14)	1.56	0.19 (2)	1.66
Kidney	3.82 (40)	3.79	2.80 (30)	2.34
Leukaemia	2.10 (22)	3.25	2.43 (26)	2.21
Lymphoma	2.29 (24)	4.94	3.08 (33)	4.07
Myeloma	2.58 (27)	1.72	1.68 (18)	1.41
Oesophagus	3.24 (34)	3.27	1.31 (14)	1.63
Pancreas	4.01 (42)	2.70	4.48 (48)	2.75
Stomach	1.81 (19)	2.32	2.15 (23)	1.26
Testis	0.19 (2)	1.34	-	-
Uterus	-	-	3.64 (39)	5.27
Brain	1.43 (15)	1.59	1.12 (12)	1.23
Colorectal	12.69 (133)	12.46	12.98 (139)	10.43
Lung	17.56 (184)	13.34	14.75 (158)	11.86
Ovary			4.20 (45)	4.32
Oral	1.34 (14)	3.07	0.75 (8)	1.64
Melanoma	1.53 (16)	4.30	2.24 (24)	4.45
Prostate	14.98 (157)	26.35		
Other	25.0 (262)	9.64	23.72 (254)	8.58
Total	100 (1074)	100	100 (1071)	100

3.5.10. **Symptoms**

Table 25 shows the most frequently occurring symptoms in the 28 days before the index date, and cancer incidence in patients with normal and raised inflammatory markers with these symptoms. None of the symptoms identified are high-risk symptoms warranting urgent cancer referrals under current NICE guidelines. The commonest symptoms were abdominal symptoms, joint

symptoms, infective symptoms, and non-specific symptoms. Cancer incidence was significantly higher for those with a raised versus normal inflammatory markers in all symptom subgroups except patients with throat symptoms. Positive predictive values were >5% for those with raised inflammatory markers associated with cough, back pain, nausea and vomiting, and chest pain.

Table 25: *Top 20 most frequently occurring symptoms in the 28 days before the index date, and cancer incidence in patients with normal and raised inflammatory markers with these symptoms**

	Normal inflammatory marker cancer incidence % (95% CI)	**Raised inflammatory markers cancer incidence % (95% CI)**	**p-value**
Abdominal pain (n=10,011)	1.49 (1.20-1.77)	4.97 (4.17-5.77)	<0.001
Tiredness (n=8,332)	1.10 (0.85-1.36)	2.84 (2.09-3.59)	<0.001
Cough (n=5,801)	2.03 (1.54-2.51)	5.29 (4.33 – 6.25)	<0.001
Joint pain (n=3,678)	1.31 (0.87 – 1.75)	2.34 (1.45 – 3.24)	0.023
Diarrhoea (n=3,463)	1.01 (0.60 – 1.42)	3.11 (2.09 – 4.12)	<0.001
Headache (n=3,451)	0.96 (0.58 – 1.35)	2.07 (1.12 – 3.02)	0.012
Back pain (n=3,141)	1.81 (1.24 – 2.37)	6.21 (4.57 – 7.84)	<0.001
Lower back pain (n=2,937)	1.81 (1.23 – 2.38)	5.32 (3.71 – 6.93)	<0.001
Dizziness (n=2,854)	1.58 (1.04 – 2.11)	3.85 (2.37 – 5.34)	<0.001
Pain generalised (n=2,832)	1.19 (0.70 – 1.69)	3.42 (2.23 – 4.60)	<0.001
Chest pain (n=2,700)	1.83 (1.22 – 2.44)	6.13 (4.41 – 7.84)	<0.001
Knee pain (n=2,597)	0.56 (0.20-0.93)	2.42 (1.37 – 3.47)	<0.001
Rash (n=2,456)	0.86 (0.41-1.32)	2.84 (1.61-4.08)	<0.001
Urinary tract infection (n=2,334)	2.27 (1.44-3.11)	4.56 (3.21-5.90)	0.003
Nausea & vomiting (n=1,812)	1.66 (0.92-2.40)	5.47 (3.67-7.26)	<0.001
Shoulder pain (n=1,803)	0.99 (0.41-1.57)	4.70 (4.07-15.7)	<0.001

Malaise (n=1,691)	1.11 (0.45-1.76)	3.43 (2.05-4.82)	0.001
Throat symptoms (n=1,660)	0.79 (0.24-1.33)	1.64 (0.57-2.71)	0.12
Chest infection (n=1,579)	1.82 (0.84-2.81)	4.61 (3.09-6.12)	0.003
Low mood (n=1,415)	0.94 (0.32-1.55)	3.13 (1.30-4.95)	0.004

**Frequency of symptoms in untested patients was too low to allow calculations of cancer incidence in symptomatic untested subgroups*

3.5.11. **Sensitivity analyses**

In order to check the robustness of the data and explore potential sources of bias a number of sensitivity analyses were done (**Table 26**). Firstly, analysis was restricted to those with ≥1 year of follow up, to ensure that loss to follow up was not introducing significant bias. Secondly, analysis was repeated using the laboratory specified upper limit of normal to define a raised inflammatory marker (rather than the mean upper limit of normal see **2.4.4** methods chapter). Thirdly analysis was restricted to only the 110,245 patients eligible for Cancer Registry linkage. Fourthly, analysis was restricted to patients without pre-existing infections or autoimmune conditions (in other words keeping the same cohort as defined in section 3.3.1). Finally, patients with myeloma were excluded; on the basis that the association between raised inflammatory markers and myeloma was well established, and to demonstrate that our association persisted when these patients were discounted.

These sensitivity analyses gave minor differences in results which were not clinically significant. Restricting analysis to the 110 245 patients eligible for Cancer Registry linkage increased cancer incidence marginally to 3.82% (3.58 to 4.05) in the raised inflammatory marker group, 1.63% (1.53 to 1.73) in the normal inflammatory marker group and 1.04% (0.90 to 1.17) in the untested group.

Table 26: *Sensitivity analyses - cancer incidence in those with raised inflammatory markers, normal inflammatory markers and untested controls, in selected subgroups*

Subgroup	**Cancer incidence % (95% CI)**		
	Raised inflammatory markers	**Normal inflammatory markers**	**Untested**
Full cohort (n=194,777)	3.53 (3.37-3.70)	1.50 (1.43-1.58)	0.97 (0.87-1.07)
Restricting to patients with ≥1 year follow up in CPRD (n=158,025)	3.87 (3.67-4.06)	1.60 (1.52-1.68)	1.02 (0.91-1.13)
Using laboratory specified upper limit of normal (n=194,777)	3.38 (3.22-3.54)	1.50 (1.43-1.58)	0.97 (0.87-1.07)
Restricting to patients eligible for cancer registry linkage (n=110,245)	3.82 (3.58-4.05)	1.63 (1.53-1.73)	1.04 (0.90-1.17)
Restricting to patients without pre-existing autoimmune disease or infection(194) (n=174,500)	3.70 (3.51-3.89)	1.52 (1.44-1.60)	0.94 (0.85-1.04)
Excluding patients with myeloma diagnosis (n=194,698)	3.45 (3.28-3.62)	1.48 (1.41-1.56)	0.95 (0.86-1.05)

3.6. Objective 5: Inflammatory markers and mortality

Objective 5 was to explore the association between inflammatory markers and one-year mortality in primary care. This objective was further broken down as follows:

5a) What is the predictive value of inflammatory markers, singly and in combination for one-year mortality? This is addressed in section **3.6.2** to **3.6.6**

5b) How does this vary by age and gender? This is addressed in section **3.6.2**

5c) What is the association between inflammatory marker level and mortality? This is addressed in section **3.6.7**

5d) What is the cause of death in patients with a raised inflammatory marker? This is addressed in section **3.6.8**

The results described below have been published in the following Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license:

- Watson J, Whiting P, Salisbury C, Banks J, & Hamilton W. (2020). Raised inflammatory markers as a predictor of one-year mortality: A cohort study using primary care electronic health record data. BMJ Open. https://doi.org/10.1136/bmjopen-2019-036027

3.6.1. Study sample

For this analysis patients were excluded from the analysis only if the inflammatory marker test result was missing (n=673) or if results were so abnormal as to be considered spurious (n=2), to avoid excluding any patients on the basis of multimorbidity. This left a total of 159,325 patients with inflammatory marker tests and 39,928 age, sex and practice matched controls

without inflammatory marker testing. Linkage to ONS death registry data was available for 109,966. Cause of death was available from death certification data in 3,141 out of 5,512 deaths where ONS linkage was available.

3.6.2. Overall mortality by test results

In total 5,512 patients died within one-year of the index date: 648 deaths in the untested group, 1,572 deaths in the normal inflammatory marker group and 3,292 deaths in the group with one or more raised inflammatory marker.

Table 27 shows one-year mortality subdivided by age, gender and inflammatory marker test result. Patients with a raised inflammatory marker (n=47,797) had an overall one-year mortality of 6.89%, compared to 1.41% in those with normal inflammatory markers ($p<0.001$). In the untested comparison cohort, one-year mortality was 1.62%. The association between raised inflammatory markers and one-year mortality was seen in all age groups apart from the under 30-year-olds. In older age groups the absolute increase in risk was considerable; a raised inflammatory marker in the over 80s was associated with a one-year mortality of 21.8%, compared to 8.6% in the over 80s with normal inflammatory markers.

Men with a raised inflammatory marker had a significantly higher one-year mortality rate than women with a raised inflammatory marker (9.78% versus 5.29%). Patients with a raised CRP had a one-year mortality of 8.76% compared to 4.99% for those with raised ESR and 4.66% for raised PV.

***Table 27**: One-year mortality (%, 95%CI) subdivided by age, gender and test result*

	Untested (n=39,928)	**Normal inflammatory markers (n=111,528)**	**Any raised inflammatory marker (n=47,797)**	**Raised CRP (n=29,164)**	**Raised ESR (n=23,138)**	**Raised PV (n=4,568)**
Overall (n=199,253)	1.62 (1.50-1.75)	1.41 (1.34-1.48)	6.89 (6.66-7.11)	8.76 (8.43-9.08)	4.99 (4.71-5.27)	4.66 (4.05-5.27)
Age						
<30 (n=21,732)	0.09 (0.02-0.18)	0.04 (0.01-0.08)	0.08 (0.00-0.17)	0.13 (0.00-0.27)	0.00 (0.00-0.00)	0.0 (0.00-0.00)
30-39 (n=22,718)	0.04 (0.00-0.10)	0.08 (0.03-0.13)	0.37 (0.19-5.47)	0.50 (0.23-0.77)	0.28 (0.06-0.51)	0.0 (0.00-0.00)
40-49 (n=31,588)	0.19 (0.08-0.30)	0.15 (0.09-0.20)	1.12 (0.86-1.39)	1.54 (1.14-1.95)	0.92 (0.59-1.25)	0.90 (0.01-1.69)
50-59 (n=35,044)	0.41 (0.26-0.56)	0.41 (0.32-0.49)	2.37 (2.03-2.70)	2.98 (2.48-3.47)	1.79 (1.38-2.21)	1.62 (0.75-2.50)
60-69 (n=35,094)	0.84 (0.62-1.05)	0.91 (0.78-1.05)	4.96 (4.51-5.40)	6.72 (6.05-7.39)	3.88 (3.33-4.43)	3.27 (2.12-4.42)
70-79 (30,251)	2.27 (1.90-2.64)	2.51 (2.26-2.76)	9.39 (8.77-10.0)	11.4 (10.5-12.2)	7.35 (6.54-8.16)	6.32 (4.75-7.88)
>80 (22,826)	8.88 (8.05-9.70)	8.61 (8.07-9.16)	21.8 (20.9-22.7)	25.9 (24.7-27.1)	16.1 (14.9-17.3)	15.6 (12.9-18.3)
Gender						
Male (n=75,787)	1.86 (1.64-2.07)	1.58 (1.46-1.70)	9.78 (9.33-10.2)	11.5 (10.9-12.1)	7.98 (7.35-8.61)	6.63 (5.44-7.81)
Female (n=123,466)	1.48 (1.33-1.63)	1.30 (1.22-1.39)	5.29 (5.04-5.54)	6.99 (6.61-7.36)	3.65 3.36-3.94)	3.51 (2.84-4.18)

Figure 19 shows graphically one-year mortality stratified by age, gender and inflammatory marker test result. This shows that those with normal inflammatory marker test results are not at an increased mortality risk compared to untested controls. This is in contrast with cancer outcomes, where patients with normal inflammatory markers have an increased cancer risk compared to untested controls (see section **3.5.7)**.

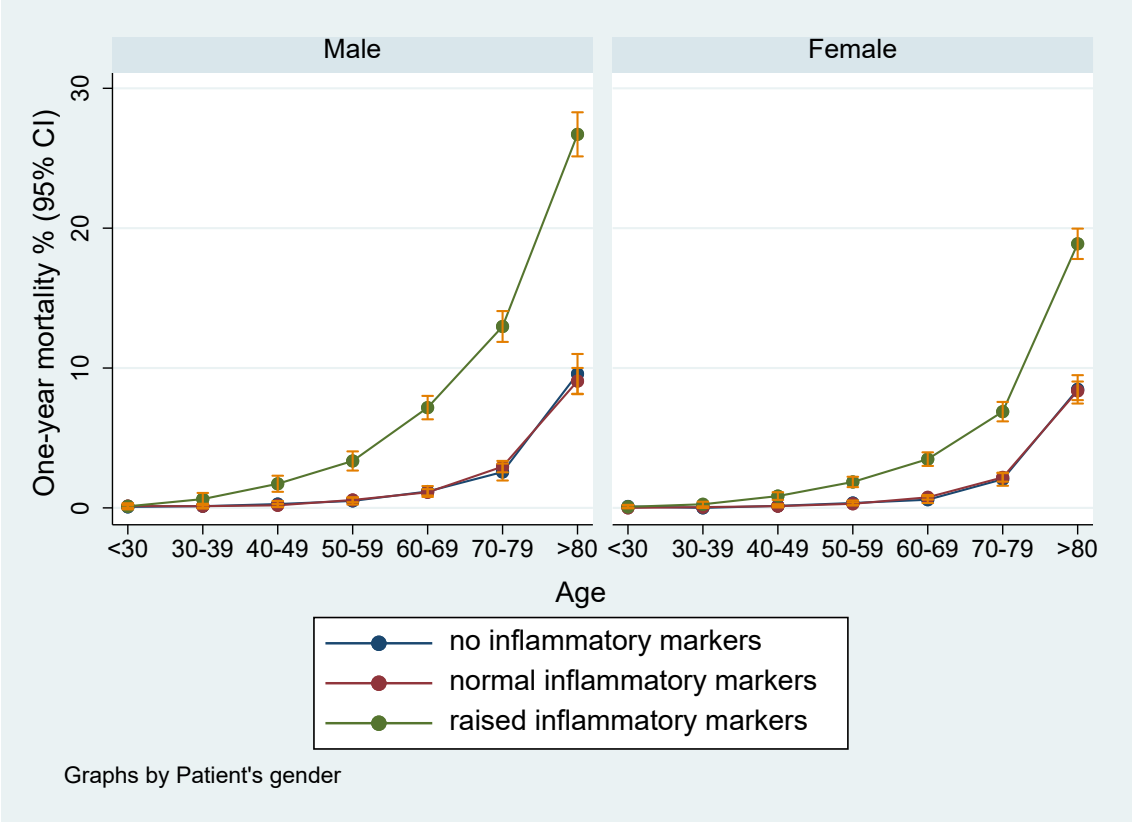


Figure 19: *One-year mortality stratified by age, gender and inflammatory marker test result*

3.6.3. Multiple inflammatory marker tests and mortality

In the 62,789 patients with more than one inflammatory marker performed simultaneously on the index date, one-year mortality was higher in the 9,029 patients with concordant raised values at 6.94% (95% CI 6.42 to 7.47), compared to the 13,783 with discordant results (one raised, one normal) who had a one-year mortality of 2.77% (2.50-3.04). In the 39,977 patients with two simultaneous negative inflammatory markers one-year mortality was 0.85% (0.76 to 0.94).

3.6.4. Mortality and repeat inflammatory marker testing

Figure 20 shows the one-year mortality in patients according to the subsequent repeat inflammatory marker results, using the most common test performed, CRP. The fact that a CRP test was requested by a GP was in itself, predictive of increased mortality, with one-year mortality of 3.30% in the tested versus 1.62% in the untested cohort. This increased to 8.76% one-year mortality if a single CRP test was raised, 9.13% if a second test was persistently raised and 14.5% the second test was raised further still. Those with a raised inflammatory marker which was not subsequently rechecked had a one-year mortality rate of 10.2%, compared to 3.25% if a subsequent CRP normalised.

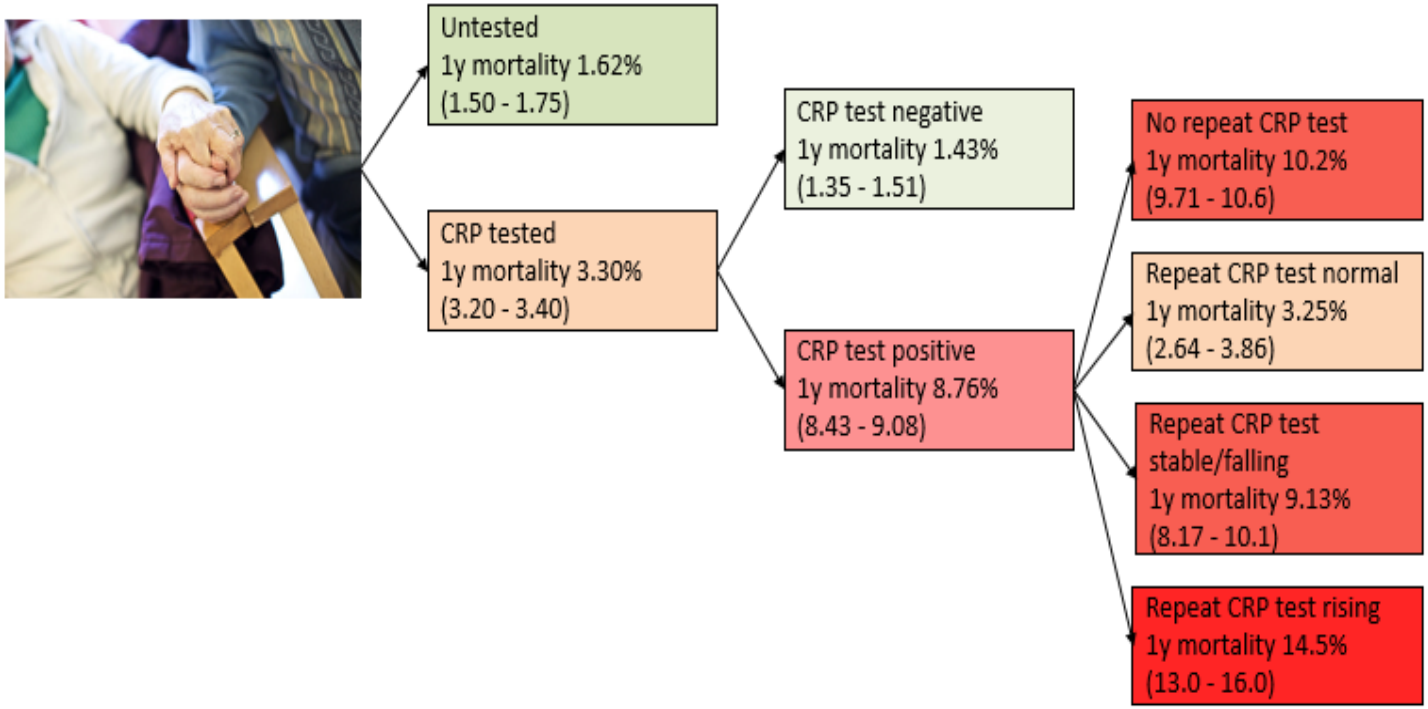


Figure 20: *Flowchart of one-year mortality (95% confidence intervals) according to CRP test results.*

The right-hand column shows one-year mortality according to repeat test result; defined as the first CRP test performed in the 3 months following the index date.

3.6.5. Measures of diagnostic accuracy - sensitivity and specificity

Table 28 shows the performance characteristics of inflammatory markers, including sensitivity, specificity and AUC. CRP had the highest sensitivity of the three tests at 67.8% and the greatest AUC at 0.78. Odds ratios (OR) reduced after adjustment for age and gender but were still significant with an adjusted OR for a raised CRP of 4.5 ($p<0.001$), 2.9 for raised ESR and 2.1 for raised PV.

Table 28: *Performance characteristics of CRP, ESR and PV for predicting one-year mortality*

						Univariable logistic regression		**Adjusted for age and gender**	
	Sensitivity	**Specificity**	**PPV**	**NPV**	**AUC***	**OR**	**P**	**OR**	**P**
CRP	67.8% (66.3-69.3)	75.9% (75.6-76.2)	8.76 (8.43-9.08)	98.6 (98.5-98.7)	0.78 (0.77-0.78)	6.63 (6.18-7.11)	<0.001	4.49 (4.18-4.83)	<0.001
ESR	56.6% (54.4-58.7)	75.7% (75.4-75.9)	4.99 (4.71-5.27)	98.7 (98.6-98.8)	0.66 (0.65-0.67	4.05 (3.70-4.42)	<0.001	2.92 (2.66-3.20)	<0.001
PV	52.0% (47.0-56.9)	72.1% (71.3-72.8)	4.66 (4.05-5.27)	98.3 (98.0-98.5)	0.62 (0.60-0.64)	2.79 (2.29-3.39)	<0.001	2.12 (1.73-2.61)	<0.001

**AUC calculated using log transformed test results as a continuous variable*

3.6.6. Measures of diagnostic accuracy – area under curve (AUC)

The logistic regression model containing age (as a continuous variable) and gender, had an AUC of 0.85, compared to 0.89 for a full model containing age, gender and CRP test result as a continuous variable ($p < 0.001$); 0.88 with age, gender and ESR ($p < 0.001$); and 0.87 with age gender and PV ($p < 0.001$).

3.6.7. Mortality and inflammatory marker levels

A dose response relationship was found between CRP result as a continuous variable and one-year mortality (**Figure 21**). In 2,184 people with a CRP >=100mg/L overall one-year mortality was 20.2%. Similar associations, with wider confidence intervals, were found for ESR and PV (not shown).

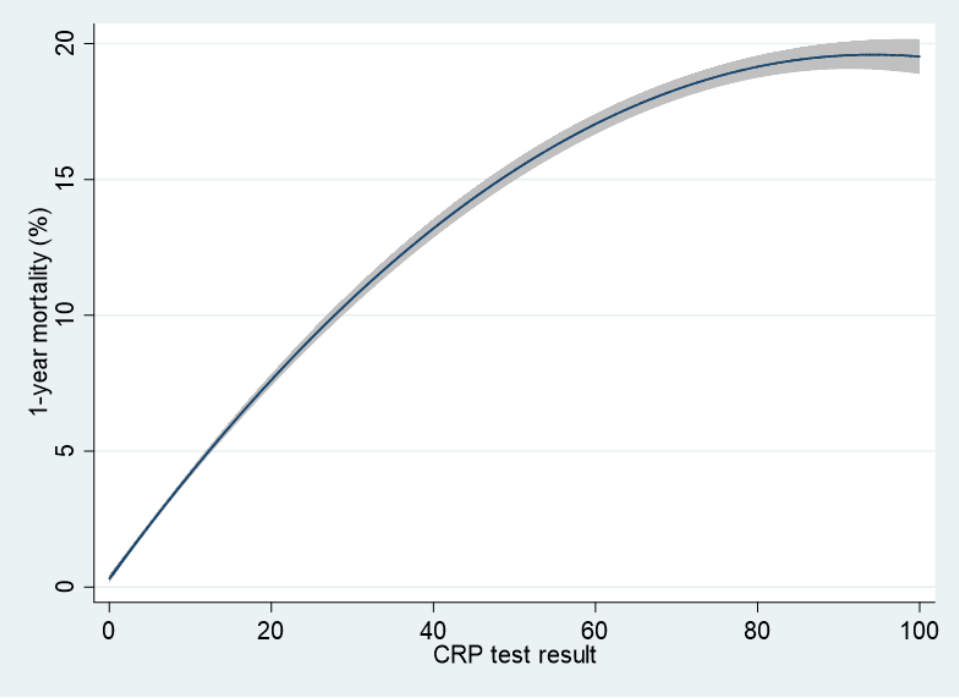


Figure 21: *Polynomial logistic regression of mortality against CRP test result as a continuous variable*

3.6.8. Cause of death

Cause of death from ONS death certification was available for 3,141 out of 5,512 total deaths in the cohort. **Table 29** summarises the cause of death amongst patients with raised inflammatory markers, compared to those with normal inflammatory markers, and untested controls. The commonest cause of death in the 26,507 patients with raised inflammatory markers was cancer (696 deaths) followed by cardiovascular disease (449 deaths). Odds of mortality in the raised versus normal inflammatory marker groups was highest for cancer (adjusted OR 6.34), followed by infections (adjusted OR 4.11). However significant increased odds of mortality were seen for all disease categories with the exception of deaths due to falls, musculoskeletal causes and senility.

Table 30 shows cause of death stratified by age group for patients with raised inflammatory markers; cancer was the commonest cause of death in 40 to 79 year-olds, cardiovascular disease increased with age and was the commonest cause of death in the over 80 age group.

Table 29: *Cause of death amongst patients with ONS death registry linkage (n=109,966)*

Cause of death	Untested controls (n=22,069)		Normal inflammatory markers (n=61,390)		Raised inflammatory markers (n=26,507)		Comparison between normal and raised inflammatory markers	
	Number of deaths	1 year mortality (%)	Number of deaths	1 year mortality (%)	Number of deaths	1 year mortality (%)	Unadjusted odds ratio (CI)	Odds ratio, adjusted for age and gender (CI)
All-cause mortality	380	1.74	889	1.45	1,872	7.08	5.16* (4.76-5.60)	3.66* (3.37-3.99)
Cancer	86	0.39	195	0.32	696	2.63	8.46* (7.21-9.93)	6.34* (5.40-7.46)
Cardiovascular disease	115	0.52	295	0.48	449	1.69	3.57* (3.08-4.14)	2.18* (1.87-2.54)
Respiratory	53	0.24	141	0.23	264	1.00	4.37* (3.56-5.36)	2.68* (2.18-3.30)
Dementia	38	0.17	69	0.11	119	0.45	4.01* (2.98-5.39)	2.21* (1.64-2.99)
Gastrointestinal	16	0.07	42	0.07	86	0.32	4.75* (3.29-6.88)	3.58* (2.46-5.20)
Genitourinary	8	0.04	16	0.03	39	0.15	5.65* (3.16-10.1)	3.13* (1.74-5.64)
Infection	5	0.02	11	0.02	29	0.11	6.11* (3.05-12.2)	4.08* (2.03-8.23)
Blood disorder	2	0.01	11	0.02	25	0.09	5.27* (2.59-10.7)	3.11 (1.52-6.37)
Senility	12	0.05	8	0.01	21	0.08	6.08* (2.69-13.7)	2.80 (1.23-6.38)
Musculoskeletal	3	0.01	8	0.01	14	0.05	4.05* (1.70-9.67)	2.40 (1.00-5.77)
Falls	4	0.02	10	0.02	6	0.02	1.39 (0.51-3.82)	0.80 (0.29-2.22)
Other	38	0.17	83	0.14	124	0.47	3.47* (2.62-4.59)	2.41* (1.82-3.20)

**p<0.001*

Table 30: *Cause of death amongst patients with one or more raised inflammatory markers at the index date subdivided by age category (n=1,872 deaths)*

Age group	Cause of death n (%)			
	Cancer	Cardiovascular	Respiratory	Other
<30	0 (0)	0 (0)	0 (0)	1 (100)
30-39	4 (40.0)	1 (10.0)	0 (0)	5 (50.0)
40-49	21 (45.7)	4 (8.7)	1 (2.2)	20 (43.5)
50-59	56 (56.6)	16 (16.2)	9 (9.1)	18 (18.2)
60-69	152 (61.0)	31 (12.5)	25 (10.0)	41 (16.5)
70-79	199 (46.6)	88 (20.6)	59 (13.8)	81 (19.0)
≥80	264 (25.4)	309 (29.7)	170 (16.4)	297 (28.6)
Total	696 (37.2)	449 (24.0)	264 (14.1)	463 (24.7)

3.7. Chapter summary

This chapter describes the diagnostic utility of inflammatory markers (CRP, ESR and PV) in primary care for relevant disease, namely infection, autoimmune disease and cancer. I will summarise the findings in relation to each of the objectives listed in **1.9.2**.

3.7.1. Objective 1: Epidemiology of inflammatory marker testing

Testing multiple simultaneous inflammatory markers was common, and abnormal results were frequent, particularly in older age-groups. Testing was more common in women, in white ethnic groups and in more affluent socioeconomic groups. Conversely, abnormal results were more common in patients from the most socially deprived socioeconomic groups. This is in keeping with the inverse care law; with potential over-testing in the affluent, and relative undertesting in more deprived groups. Higher testing rates may also in part reflect higher consultation rates in certain socio-demographic groups. Repeat inflammatory marker testing is common, particularly in those with an

initial raised inflammatory marker; most repeat tests were done within 3 months of the index date.

3.7.2. Objective 2: Inflammatory markers and overall disease outcomes

In patients with a raised inflammatory marker, the most common diagnoses were infection (6.3%), followed by autoimmune conditions (5.6%) and cancers (3.7%). Disease incidence was higher in those with more than one simultaneous raised inflammatory marker (22.6%) and in those with a second repeat test which was persistently raised (23.8%).

Inflammatory markers have low sensitivity (<50%) and are therefore not useful as a 'rule-out' test. Instead, they are classic Bayesian tests, with a positive test somewhat increasing the chance of disease whilst not being definitive, and a negative test reducing the chance of disease, again not to zero.

Disease incidence increases with rising inflammatory marker levels in a dose-response relationship.

For any relevant disease, small differences were seen between the three tests: areas under receiver operating curve (AUC) ranging from 0.659–0.682. CRP has the highest overall AUC, largely because of marginally superior performance in diagnosis of infection (AUC CRP 0.617 versus ESR 0.589, $p<0.001$). The three tests are equivalent for diagnosis of autoimmune diseases and cancers.

Adding a second test gave limited improvement in the AUC for relevant disease (CRP 0.682 versus CRP+ESR 0.688, $p<0.001$); this is unlikely to be of clinical value, even if statistically significant. The negative predictive value for any single normal inflammatory marker was 94.0%, compared to 94.1% with multiple normal tests. No combination of inflammatory marker tests can be used to rule-in or rule-out disease confidently. The maximum sensitivity of 60.6% (for the

combined test CRP|PV) is low, yet comes at a price of increased false positives compared to using single tests.

3.7.3. Objective 3: Symptoms and cascade testing

Although the CPRD data does not give information on the reasons for inflammatory marker testing, the frequency of non-specific symptoms such as tiredness suggests that these tests are commonly used as a non-specific test for underlying systemic disease.

Raised inflammatory markers are associated with higher rates of GP consultations, blood tests, and referrals compared to normal inflammatory markers. This is likely to relate to the wide range of potential differential diagnoses in patients with raised inflammatory markers. Although the unit cost of inflammatory marker tests is relatively low, the total costs, including these follow-on consultations, investigations and referrals are likely to be substantial.

3.7.4. Objective 4: Inflammatory markers and cancer

Primary care patients with a raised inflammatory marker have an overall one-year cancer incidence of 3.53%, more than twice the risk in those with a normal test. Cancer incidence rises with rising levels of inflammatory markers and is higher still if a second test shows persistent raised inflammatory markers. However, inflammatory markers are not a useful rule-out test for cancer, as with a sensitivity of 46.1% for CRP, 43.6% for ESR and 49.7% for PV, roughly half of the tested patients with subsequent cancer diagnosis had a normal inflammatory marker test. Patients with normal inflammatory markers have a cancer incidence of 1.50%, higher than the untested group with 0.97% cancer incidence. Importantly, women under 60 and men under 50 with raised inflammatory markers have a risk of cancer below 3% so investigations for cancers would not usually be warranted.

3.7.5. Objective 5: Inflammatory markers and mortality

Inflammatory markers are a strong predictor of all-cause mortality in primary care. The association between raised inflammatory markers and all-cause mortality is seen in all age groups except patients aged less than thirty years. Men with raised inflammatory markers have a higher one-year mortality than women (9.78% versus 5.29%). Of the three tests examined, CRP has the highest predictive accuracy for mortality.

Further discussion, including the clinical implications and strengths and limitations of these findings can be found in **Chapter 6**; the overall discussion and conclusions chapter.

CHAPTER 4. QUALITATIVE METHODS

4.1. Chapter overview

This chapter describes the methods of the qualitative study. The overall aim is to explore the meaning of inflammatory marker tests for doctors and patients in primary care. Whilst the quantitative study provided information to improve clinicians' understanding of the diagnostic utility of inflammatory markers in primary care, this is not sufficient to improve clinical practice; doctors need to share this information about inflammatory marker testing appropriately with patients.

The specific objectives of this study were:

1) To explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests.
2) To provide in-depth exploration of patients' experiences of testing – from GP consultation to results.
3) To identify barriers and facilitators to improved communication, in order to inform improved communication in future.

Qualitative methods were chosen as the most suitable way of providing in-depth exploration of doctors' and patients' experiences. To achieve these objectives, I used qualitative interviews with patients who recently had inflammatory marker blood tests in primary care, and the GPs who requested these tests. To achieve objective 1, it was important for me to obtain paired data to allow me to compare doctors' and patients' differing perspectives on the same healthcare encounter. To achieve objective 2, it was important for me to interview patients both at the time of testing, and also after they had received their blood test results.

In this chapter I first describe the methods of sampling and recruitment, data collection and analysis. Next, I explore reflexivity and the ethical issues raised by the study. Finally, I discuss the patient and public involvement which underpinned the research.

4.2. Sampling and Recruitment

Practices were recruited via the West of England Primary Care Clinical Research Network eBulletin. Practices were purposively selected to include a range of urban and rural practices, and to reflect a range of practice populations in terms of deprivation, age, and ethnicity. Participating practices were visited prior to commencing recruitment and the study was explained to GPs at a practice meeting. Participant Information Sheets (PIL) were then emailed to all GPs via the practice manager. All GPs in participating practices were eligible to participate, including locums, salaried GPs, and partners. GPs who were interested in participating were asked to complete a contact sheet which enabled them to be contacted once a patient whom they had consulted had been recruited.

Patients were eligible to participate if they were:

- aged >18 years.
- having blood tests which included CRP, ESR or plasma viscosity (henceforth referred to as 'blood tests').
- having tests requested by participating GPs (who had completed a contact sheet).
- able to speak English sufficiently for interview.

The sampling strategy for patients was purposive, in order to include a range of patient gender, age and socioeconomic status. Eligible patients were provided

with information about the study by their healthcare team but were under no pressure to participate.

Methods of patient recruitment were flexible, to accommodate different practice procedures for taking blood for testing. In five out of six practices, I was able to sit in a spare room during phlebotomy clinics. At the start of these clinics, I met with the phlebotomists to explain the study and eligibility criteria. Phlebotomists then offered eligible patients a Participant Information Leaflet (PIL). If participants were interested, they were invited to discuss the study with me. I could then either complete the interview immediately or arrange a suitable interview time at a location of the patient's choice. Two practices supplemented this by GPs explaining the study at the time of requesting blood tests and obtaining permission for the research team to contact patients. One practice did not have available space for a researcher to sit in surgery; in this practice all participants were recruited by phlebotomists, patient contact details were sent securely to me using an nhs.net secure email, and patient interviews were arranged at the patient's convenience.

At the end of the first patient interview, a second follow-on interview with the patient was arranged 1-2 weeks later, in order to explore the communication of the test results.

Once a patient had been recruited, the GP who had requested the blood tests was contacted to arrange a telephone interview. Each GP could complete a maximum of two interviews each (about different patients), to ensure that a range of GPs were sampled. All GPs had received the test results at the time of interviewing; this is likely to reflect the fact that electronic test reporting means most results are returned to GPs within 24 hours.

For each test requested I therefore aimed to complete three interviews (henceforth referred to as 'cases'); firstly, with the patient around the time of testing, secondly with the patient after the test results were obtained, and finally with the GP who requested the test.

4.3. **Interviews**

Twenty-one of the initial patient interviews took place in participants' GP practices, six took place at the University of Bristol and one was done via telephone. All second interviews and all GP interviews were done by telephone. Interviews were semi-structured, using topic guides based on the research questions but flexible enough to allow exploration of issues raised by the participant. The topic guide (see **Appendix C**) was adapted iteratively during the study, using information emerging in early interviews to inform further questions for exploration in subsequent interviews. The first patient interview focussed on patients' experience of testing, their understanding of the rationale for testing, their expectations of testing and the communication they had received about their tests. The second patient interview allowed me to explore whether patients' views and understanding of testing changed between initial consultation and results, with questions focussing on patients' experience of receiving and interpreting their test results. The GP interviews allowed me to compare patient and GP perspectives of reasons for testing, expectations of tests and communication around testing. GPs undertook the interviews with access to the patient's electronic medical records at the time of interviewing as an aide memoire.

The interviews were recorded, with consent, using a digital encrypted audio recorder. Written informed consent was obtained at the start of each interview; for telephone interviews verbal consent was recorded and documented replicating the consenting text used on the written consent forms.

Patients received £20 high street voucher payment to recompense their time participating. GP practices were paid £60 for each GP interview completed, in keeping with Primary Care Research Network agreed rates. Interviews were continued until data saturation was achieved, meaning that a diverse sample had been achieved, the topic guide was stable and no new codes were arising in the data.(206)

4.4. Fieldnotes

Fieldnotes were recorded during and immediately after each encounter. These included observations, reflections and evolving analytical thoughts. These were used as a prompt and reminder prior to follow up interviews with patients. They were also used to aid the iterative development of the topic guide, including changes and additions to the flow and structure of the questions and prompts. Finally, they were used when re-reading and listening to the audio-recorded interviews to help with recollection of the non-verbal communication and interactions, which were recorded in NVivo as memos, and contributed to the evolving analytical ideas. Examples of the use of these field notes included a patient who was reticent and appeared anxious during the initial interview. This anxiety was not verbally expressed, and the patient was reluctant to open up, so this was not easy to recognise in the first interview transcript. Memos were added to NVivo to reflect my observations of the non-verbal communication, and my field notes were re-read ahead of the second interview. I used careful probes in the second interview to help improve the flow of the conversation and develop rapport and managed to get a better exploration of the patient's experiences and anxieties in the second interview.

4.5. Analysis

Audio recordings were transferred to secure servers at the University of Bristol and transcribed verbatim by an experienced University of Bristol approved transcriber. I listened to all audio-recordings and cross-checked these with the transcripts making minor corrections. All written data were anonymised. Analysis began when the first transcripts were available, so that data collection and analysis were conducted concurrently, with early interviews informing the

iterative development of the topic guide for subsequent interviews. This process of iterative development did not continue throughout, the topic guide stabilised as saturation was achieved. Data were read and re-read to aid familiarity.

All transcripts were uploaded into NVivo and files were classified into three categories: initial patient interviews, follow up patient interviews and GP interviews. Files were also organised into cases which included data from the same patient before and after testing and the paired GP interviews.

I analysed the transcripts using thematic analysis, involving a mixture of inductive and deductive coding and constant comparison.(207) This involved an iterative process of close reading of the data, coding, constant comparison and elaboration of emerging themes. Inductive coding aims to discover the meaning of the data without making assumptions based on previous understanding; by comparison deductive coding is informed by existing theories and by the research questions. It would not be realistic to claim that I could (or should) completely free myself from my own preconceptions given my clinical and research experience in this topic area; I was also influenced in my coding by my research questions. I therefore chose to combine a mixture of inductive and deductive approaches. Thematic analysis was chosen as this is a flexible approach, allowing me to make comparison within cases (i.e. before-after test results and between doctor-patients) and also between cases (i.e. comparing patients and GPs as a group).

My qualitative supervisor (JB) and I independently reviewed four of the transcripts (including first and second patient interviews and GP interview) to develop an initial coding framework that reflected the research objectives. This was amended and adapted iteratively following feedback and discussion with my supervisor (JB), and also following discussion with my PPI panel. The modified framework was then tested independently by myself and my supervisor (JB) using a further three transcripts. Once the final coding framework was agreed (see **Appendix D**), I then took responsibility for ongoing coding and categorisation of the data, using the NVivo qualitative data management

software. To assure the reliability of the coding and analysis process, codes and categories were reviewed in regular meetings with JB to ensure the accuracy of interpretation and internal consistency of codes.

As I developed the coding framework, I read the data case by case, and wrote a narrative case summary for each participant, to retain my understanding of each case as a whole. I identified key differences between cases, which were important in order to retain an overview of the extent of information sharing and shared decision-making in the consultations. I chose to further analyse the data using qualitative content analysis to categorise and capture this information.(208) I used a coding framework which had been developed and tested for the purpose of analysing and quantifying the extent of shared decision-making in video-recorded primary care consultations by a Foundation Year 2 doctor Jess Martin whom I supervised during the course of my PhD.(181) I made minor adaptions to this framework in response to the data in my interviews. This allowed me to categorise each case based on whether the patient knew which tests were being done, why tests were being done, whether decision-making was shared or not, how results were communicated and whether they understood the meaning of their test results. This provided an overview of my data, which triangulated and strengthened the findings of my thematic analysis. These categorical codes were recorded using NVivo case classifications; definitions of each code are detailed in Chapter 5 (**Table 35**, **Table 36** and **Table 37**).

When all the data were coded, I began to make comparisons between doctors' and patients' interviews, scrutinising the data for areas of congruence and dissonance. As well as exploring emerging themes and patterns I sought out apparently contradictory quotes and examples; these were coded and noted as memos in NVivo. I met regularly with JB to discuss my understanding of the data and developing themes. Categories of data and thematic relationships were identified and written up as descriptive and interpretive accounts, supported by interview excerpts.

4.6. **Reflexivity**

Reflexivity was important for me to consider, as I have a dual role as both a researcher and a GP. My own experiences as a GP therefore influence my perspectives and my interactions with the research participants.

It is recognised that research participants are influenced by how they identify with a researcher.(185) I chose to be open with participants (both patients and GPs) and to identify myself as being a GP, as it felt dishonest to withhold this information. I was influenced in this decision by the discussions and advice from my Patient and Public Involvement group (see section 4.8)

In the GP interviews I emphasised that the interviews were non-judgemental, and were focussed on exploring communication around testing, not on scrutinising the clinical decision-making. During my interviews I noticed that the fact I was a doctor did impact on the way some GPs interacted with me. Some slipped into the use of medical jargon or made assumptions that I had a shared understanding of testing and medical knowledge. I therefore had to prompt to ensure they explicitly explained their rationale for testing in lay terms where necessary. In a small number of interviews, the GPs I spoke to were clearly aware of my published research into inflammatory marker testing. This may have impacted on the way they presented their diagnostic argument for testing, to try to fit with what they thought they 'should' do. However, my experience overall was that GPs were very open with me, their reflection on uncertainties around inflammatory marker testing indicated they were articulating their actions and thought processes at the time of the consultation. Overall, being a GP did not prove to be a barrier and may have facilitated my communication with GPs as they felt comfortable discussing cases with a fellow clinician with shared understanding.

I anticipated that by identifying myself as a GP, patients might ask me questions about their health. In practice this was not a significant issue. On the small number of occasions where patients asked me for my views or opinions, I felt

comfortable explaining that I had no access to their medical records or information about their tests and was not in a position as a researcher to be able to make any comments. In one case I felt slightly conflicted as the patient had not yet obtained their test result, but I knew from my interview with the GP that it had shown a significant abnormality. As I was not able to disclose this information, I recommended that they should contact the practice for their test results, which they assured me they would.

It was possible that because I identified as a GP, patients could have been reluctant to share negative information about their GP with a fellow doctor. However, my experience did not seem to match with this, as patients appeared to be very open and were often forthright in sharing their experiences, including negative experiences. The fact that I identified with participants as a GP did appear to enhance my credibility with patients and helped enable my access to the practices as a researcher. It also seemed to help some patients to feel comfortable discussing medical information with someone whom they identified as trustworthy by virtue of being a doctor.

I kept reflective notes throughout the study, identifying points where I felt the doctor or patient had been influenced by my status as a GP. During the analysis, I added an additional *post hoc* code entitled 'study impact' in which I coded any examples where a participant seemed to have been influenced by their participation in the study, so that I could take this into account during my analysis.

I was aware that my role as a clinician could affect my analysis, as I have my own experiences and preconceptions of blood testing. This was addressed by ensuring that my non-clinical qualitative supervisor (JB) read and coded the initial interviews and helped develop the coding framework and also by involving my PPI group in the early development of the coding framework. Throughout the course of my analysis, I had regular meetings with JB to check and discuss my evolving analytical thoughts with a non-clinician.

4.7. Research governance and approvals

The Integrated Research Application System (IRAS) system was used to obtain ethical approval, Health Research Authority (HRA) approval and confirmation local capacity and capability. The study was adopted and included in the NIHR Clinical Research Network (CRN) portfolio. Ethical approval was obtained from the proportionate review sub-committee of the London – Hampstead Research Ethics committee REC reference 19/LO/0405, who concluded that '*this was a well-presented study with no material ethical issues*'.

4.7.1. Ethical issues

The main ethical considerations which were listed in the IRAS (Integrated Research Application System) application were as follows:

Recruitment

Recruitment could cause inconvenience for GP practices, GPs, and patients, which was particularly important to consider given the workload pressures primary care is facing. I minimised this by making recruitment as simple as possible and by ensuring that methods of recruitment were flexible so that different practice procedures for blood testing could be accommodated. Practice staff were not required to consent patients, only to issue Participant Information Leaflets to eligible patients. I did not approach patients directly.

Consent

Eligible patients were provided with information about the study by their healthcare team but were under no pressure to participate. As a researcher I undertake regular Good Clinical Practice training including training in obtaining

informed consent; I also have relevant experience in obtaining consent from of my clinical role as a GP.

Interviews

There was a small risk that patient interviews could cause anxiety - thinking and talking about blood tests could potentially raise questions or worries patients might not otherwise have considered. To minimise this, I piloted the topic guide with my Patient and Public Involvement (PPI) panel and adjusted it based on their feedback. My experience as a GP ensured that I was used to discussing sensitive information with patients. Although interviews did not directly discuss patients' medical history, there was a risk that patients could make medical disclosures. I emphasised that I could not provide medical advice in these circumstances and advised them to see their GP with any medical queries. I also explained that although interviews were confidential this is not absolute and disclosure would be necessary in line with my duty of care, if a safeguarding concern was identified. Although I considered these possibilities, in reality I did not have to break confidentiality at any time for these reasons.

Data handling

Storage and handling of all data complied with the Data Protection Act 2018, the General Data Protection Regulation (GDPR) 2018 and University of Bristol's data protection policies. All written data was anonymised, and participants were assigned a unique identifier. I was the only person with access to identifiable participant information in a master document linking this with the participant's unique identifier.

4.8. Patient and public involvement

A Patient and Public Involvement (PPI) group comprising five participants was established in October 2018. Participants were recruited from the Health Research Panel of the NIHR Collaboration for Leadership in Applied Health Research and Care West (CLAHRC West – now 'ARC West'), with support of the CLARHC West PPI lead. The PPI group met three times face to face for two hours on 4th October 2018, 13th December 2018, and 22nd October 2019. Meetings were chaired by myself, with my qualitative supervisor JB attending to provide support and take notes. Refreshments were provided and participants were paid at NIHR rates. A final 1.5-hour meeting was held on 16th July 2020 via online teleconferencing due to the coronavirus pandemic. Further feedback was provided via email in between meetings.

In the first meeting on the 4th of October 2018, the discussions focussed on the aims and objectives of the research and the wording of the Participant Information Leaflets. Initially I had planned to focus on how GPs communicate inflammatory marker test results to patients (see **Table 31**). The PPI group discussion highlighted the importance of broadening this and exploring the patient experience of test communication from initial GP consultation through to test results. They emphasised that this should encompass not only doctor-patient communication but also communication with phlebotomists, receptionists, nurses, and allied health professionals. They suggested this could also include patient understanding of the systems and processes of testing and obtaining test results. The PPI panel also emphasised the importance of exploring not just communication but also patient understanding of why the tests were being done, and what the results would be able to tell them and their GP. This led to changes in the overall aims and objectives of the study (see **Table 31**) which had a significant impact on the final outcomes of the study.

Table 31: *Aims and objectives before and after PPI input*

Original aims and objectives	Final aims and objectives
Aims: • To help GPs discuss inflammatory marker blood tests with patients in a way they can understand.	Aims: • To explore how an understanding of inflammatory markers is shared between doctors and patients in primary care.
Specific objectives: • Do patients understand why blood tests are being done? • Do patients understand what their test results mean? • How could doctors better explain tests to patients?	Specific objectives: • To explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests. • To provide in-depth exploration of patients' experiences of testing – from GP consultation to results. • To identify barriers and facilitators to improved communication.

Another key message from the first PPI meeting was their suggestion that the term 'inflammatory markers' should be avoided in patient interviews, unless the patient themselves mentioned it first. The PPI group explained that from a lay perspective 'inflammatory markers' sounded worrying and serious, and this could cause anxiety or confusion for participants. As doctors would usually be testing inflammatory markers alongside several other blood tests, the PPI group felt it would not be meaningful to try to discuss inflammatory markers in isolation. Instead, they recommended exploring what patients understood about the different tests they were having done and why. The Participant Information Leaflet (PIL) and topic guide were therefore amended and the term 'blood test' was used throughout. These insights helped me to consider the wider context of blood test communication and ensure the findings had broad applicability. The PPI group provided further support via email after the initial face-to-face meeting, including reviewing participant information leaflets and topic guide questions to ensure that these were written in accessible language.

In the second meeting on the 13th of December 2018, the PPI group provided in-depth feedback on the interview topic guide, making suggestions to ensure that the wording was clear and easy to understand and would not provoke anxiety. They also suggested additional questions to explore patients' understanding of

the process for obtaining their test results, and patients' understanding of the meaning of their test results. In particular, the PPI group felt it was important to explore whether patients understood the implications of test results for their health, and what patients should do with the test results.

The PPI group were also involved in planning methods of recruitment. We discussed the tension between giving participants enough time to reflect on the study and capturing the interviews as soon as possible after the initial consultation. The PPI panel emphasised the importance of ensuring that the interviews were convenient for participants. The PPI panel felt that offering the option of an interview in the GP surgery at the time of testing might make the study easier and less burdensome for many patients, but that flexibility was needed to accommodate patient preference. As a result of this feedback, I decided to make the methods of recruitment flexible, incorporating the option of sitting in a side room during phlebotomy clinics, so that patient and GP practice preferences could be easily accommodated. The method of sitting in the surgery during phlebotomy clinics turned out to be the most successful method of recruitment, accounting for the majority of interviews completed.

Finally, I also discussed with the PPI panel my own reflexivity as a researcher and a clinician, as I was uncertain how I should identify and introduce myself to research participants. After reflection, the PPI group agreed it would seem dishonest not to inform research participants that I was a GP myself. This information was included in the PIL for both patients and GPs. The PPI group emphasised the important of explaining that I was conducting the interviews in my role as an independent researcher and emphasising that the results would not be fed back to their GP and would have no bearing on their care.

During the third meeting, on 22nd October 2019, early interview transcripts of the study were shared with the PPI panel members. The findings resonated with the PPI panel members, who provided feedback which was used to help with development of the coding framework. The PPI panel noted differences in the interview transcripts between patients who were active and engaged in

communication and shared decision-making compared to those who were more passive and disengaged and additional codes were added to the coding framework to capture this.

At the final PPI meeting on 16th July 2020 the early findings of both the qualitative and quantitative aspects of the thesis were shared with the PPI group. Feedback was used to explore how the qualitative and quantitative components of the thesis could be synthesised. We discussed how the uncertainty in interpreting inflammatory marker tests uncovered in the quantitative component of the thesis could be shared with patients. We also explored future ideas for improving the communication of inflammatory marker testing in future. This is discussed in more detail in the discussion chapter **section 6.6.**

4.9. Summary

This chapter sets out the methods for the qualitative study, which explored doctors' and patients' expectations and experiences of testing in primary care. For each case, three interviews were undertaken; firstly, with the patient prior to testing, secondly with the patient after the test results had been obtained, and thirdly with the requesting GP. This provided rich data allowing me to compare and contrast the accounts and experiences of patients and doctors, to explore areas of congruence and dissonance and develop an understanding of the interactions which underpin a shared understanding of testing. The results of the qualitative study are presented in the next chapter.

CHAPTER 5. QUALITATIVE RESULTS

5.1. Chapter overview

This chapter describes the results of the qualitative study. The overall aim was to explore the meaning of inflammatory marker tests for doctors and patients in primary care. The specific objectives were:

1) To explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests.
2) To provide in-depth exploration of patients' experiences of testing – from GP consultation to results.
3) To identify barriers and facilitators to improved communication, in order to inform improved communication in future.

The data presented in this chapter cuts across these three objectives, so results will not be structured according to the objectives. Instead, I will first describe the participant sample, then the results of the thematic analysis and qualitative content analysis. The results of the qualitative content analysis are presented in **Table 35**, **Table 36** and **Table 37**. All other data are derived from the thematic analysis. The thematic analysis firstly focuses on doctors' and patients' expectations of blood tests, next the process of decision-making, communication, and shared understanding of tests, and then the consequences of testing. Finally, I discuss the barriers to a shared understanding of tests which were identified by doctors and patients, relating to knowledge, attitudes, and systems of testing. To summarise, I then reflect on the results in relation to the three main objectives.

The results described in this chapter are currently in preparation for submission to a peer-reviewed journal, with a planned submission date of June 2021.

Discussion of the strengths and limitations, comparison to existing literature and implications for clinical practice are covered in the discussion Chapter 6.

5.2. Description of the sample

In total 28 patients and 19 GPs from six GP practices were recruited. Eighty interviews were carried out between 31st May 2019 and 17th March 2020; 26 GP interviews and 54 patient interviews (most patients and some GPs were interviewed twice).

5.2.1. Practice demographics

All practices in the West of England Primary Care Research Network were invited to participate by email, and expressions of interest were received from twenty-three practices. Of these, six were chosen to participate, to reflect a range of socio-demographic characteristics (see **Table 32**). In participating practices, information sheets were sent to all GPs, and completed contact sheets were received from 40 GPs in total; this enabled me to contact the relevant GP once a patient they had consulted was recruited.

Table 32: *Demographics of participating practices*

Practice name	List size*	Deprivation score*	Rural/urban	Ethnicity*
A	20,026	Third least deprived decile	Urban	3.4% mixed, 9.3% Asian, 3.0% black, 1.2% other non-white
B	32,158	Least deprived decile	Rural	1.1% Asian 1% other non-white
C	19,903	Third more deprived decile	Suburban	3.1% mixed, 4.9% Asian, 3.2% black
D	8,960	Most deprived	Urban	5.6% mixed, 6.7% Asian, 22% black 2% other non-black
E	15,625	Second least deprived	Suburban	3% mixed, 5.1% Asian, 1.5% black
F	11,280	Third more deprived decile	Urban	2% mixed, 2% Asian, 2.1% black

**Public Health England National General Practice Profiles [accessed April 2019]*

5.2.2. Patient characteristics and reasons for testing

Table 33 summarises the characteristics of the 28 patients who participated. The proportion of women recruited (64%) is in keeping with the quantitative research from CPRD which showed that 62% of inflammatory marker tested patients were female. The most common reasons for testing (based on GP interviews) were joint symptoms (n=5), suspected polymyalgia (n=3), tiredness (n=3), bowel symptoms (n=3), monitoring of pre-existing inflammatory conditions (n=3), headache (n=2) and hospital follow up (n=2). Other reasons for testing included breathlessness, weight loss, lymphadenopathy, eating disorder, pleural effusion, gynaecological symptoms, and patient requested screening (all n=1). Out of the 28 patients recruited, 21 interviews were carried out in the GP practice at the time of their phlebotomy appointment, 7 interviews were carried out soon after testing either in the GP practice or at the University of Bristol. Out of the 28 patients recruited, 23 patients had two interviews; one at around the time of testing, and a second follow up interview. At the time of the second interview, five of these patients had still not received their test results. Three of these agreed

to be interviewed a third time to discuss their experience of receiving their results, two were not planning to obtain their results, so were not interviewed again (both had bloods done for routine monitoring purposes). Five patients had only one interview; four of these had already received their results at the time of the first interview and one patient was uncontactable for follow up.

Table 33: *Demographics of participating patients (n=28)*

Characteristic	**n (%)**
Gender	
Female	18 (64%)
Male	10 (36%)
Ethnicity	
White British	23 (82%)
BAME	3 (11%)
Other non-British	2 (7%)
Age group	
18-24	8 (29%)
25-34	3 (11%)
35-44	3 (11%)
45-54	3 (11%)
55-64	3 (11%)
65-74	1 (4%)
75+	7 (25%)
Socioeconomic status (based on postcode IMD)	
1 (most deprived)	2 (7%)
2	5 (18%)
3	2 (7%)
4	4 (14%)
5	0 (0%)
6	2 (7%)
7	2 (7%)
8	3 (11%)
9	2 (7%)
10 (most affluent)	1 (4%)
Postcode unavailable	5 (18%)

5.2.3. GP demographics

A total of 26 interviews with 19 GPs who requested the blood tests were completed. I was unable to obtain an interview from the requesting GP for two of the patient interviews. The characteristics of participating GPs are summarised in **Table 34.** Out of 19 GPs recruited, 14 were female (74%). More male GPs contributed >1 interview so overall 8 interviews (31%) were with male GPs and 18 interviews (69%) were with female GPs.

Table 34: *Characteristics of participating GPs (n=19)*

Characteristic	n (%)
Gender	
Female	14 (74%)
Male	5 (26%)
Type of GP	
Partner	13 (68%)
Salaried	5 (26%)
Locum	1 (5%)
Years' experience	
0-5 years	5 (26%)
5-10 years	2 (11%)
10-20 years	8 (42%)
20+ years	4 (21%)

5.3. Expectations of testing

Doctors and patients had differing expectations of testing. Firstly, I explore patients' expectations, then compare this with doctors' expectations. Finally, I explore the paired data to identify the impact of different expectations occurring within a doctor-patient encounter.

5.3.1. **Patients' expectations**

Few patients directly requested tests, but most saw them as 'a good thing'. Patients saw blood tests as a way of moving forward with their problem, for example, by determining the cause of unexplained symptoms, and saw testing as a sign that the doctor was taking their symptoms and concerns seriously.

> *I've been happy they've done them because you feel oh that's good because he's looking out for me in some way (Patient 18, female, 76 years)*

Patients had high expectations of their tests; they hope that tests will provide answers, and therefore lead to solutions for their symptoms.

> *I was happy she was doing blood tests and extra tests because I need answers (Patient 20, male, 76 years)*

Even if solutions are not found, having greater certainty or a diagnosis for their symptoms was felt to be important.

> *I'm quite happy with that. I feel at least I'm getting somewhere. I can get the results back and I might say, you know, I've got arthritis. (laughs) I could cope with anything. (Patient 2, female, 82 years)*

Some patients said that they were hoping for a result which would provide validation that they had an underlying reason for their symptoms.

> *I think it will make me feel a bit better knowing that it's not just me like taking care of myself badly and there's actually a reason why I might feel the way I do sometimes (Patient 17, female, 18 years)*

> *You almost hope for a result that's going to come back with something that can be treated in some way (Patient 21, female, 29 years)*

Patients have limited understanding of specific details of their blood tests but perceive them as providing a powerful way of getting an accurate validation of what is happening inside their body, and their overall health status.

> *Kind of shows what's happening on the inside. Like on the outside I could be fine and kind of- but then the inside, my blood may be having some different story or knowing before I get really ill (Patient 3, female, 20 years)*

In some atypical cases, blood tests were viewed as transactional. This was particularly noted for the three patients who were having blood tests for

monitoring rather than due to symptoms. These patients had been told to get tests done by the practice and regarded these tests as being mainly for the GP and not for themselves. As a result, the patients appeared relatively passive and uninterested in results.

> *I just know I need to get this blood test so that my medication keeps going and that's it… If there's a problem somebody will ring me. If there's no problem then nobody will ring me, but actual fact I'm fine with that, I'm absolutely fine with that (Patient 10, female, 56 years)*

5.3.2. **Doctors' expectations**

Doctors tended to have lower expectations of testing, particularly if tests were being done in patients with non-specific symptoms. In these cases, a 'panel' or 'battery' of routine tests were seen as a useful 'screen' to help to rule-out serious causes of symptoms, rather than to rule-in specific diagnoses.

> *I think it was kind of the panel really would guide me in which direction we needed to move forwards (Doctor 9, male, GP partner, 0-5 years' experience)*

> *I just arranged some baseline bloods as a bit of a general screen (Doctor 25, female, locum, 0-5 years' experience)*

> *Probably most of ours have got irritable bowel but we sort of need to make sure we've excluded those other things first…yes, excluding more undiagnosed disease first before we move to anxiety or irritable bowels as a diagnosis (Doctor 15, female, GP partner, 10-20 years' experience)*

Doctors' decision-making encompassed both medical and non-medical aspects of testing. As well as diagnostic purposes, tests were also felt to be useful to manage patient anxiety and as a tool for managing patient interactions. In these cases, doctors often had an expectation that the test result was most likely to be normal.

> *To be honest I think I was probably expecting them to come back normal and just as an added tool or, yeah, an added tool to reassure the patient (doctor 12, female, salaried GP, 0-5 years' experience)*

5.3.3. Mismatch between doctors' and patients' expectation

A mismatch in expectations could be seen when analysing the paired data on patients and doctors. For example, patient 18 expected the blood tests to give answers to the cause of her symptoms, whereas her doctor expected the results to be normal.

I think I was pretty much expecting them to be normal 'cos I didn't treat her or anything at the time. (Doctor 18, male, GP partner, 0-5 years' experience)	*Well just that he would discover what it was and confirm what it was and give me some help (Patient 18, female, 76 years)*

Some doctors were aware of this mismatch in expectations, and therefore tried to discuss and share their expectations with patients to pre-empt and prepare patients for the possibility of a normal test result.

> *What I normally do say if I think they're going to be normal, I normally try and pre-empt that by saying to patients I expect they'll be normal but that will be great and reassuring. (Doctor 23, female, salaried GP, 10-20 years' experience)*

5.4. Decision-making

Most decisions to order a blood test were led by doctors. Although doctors perceived that some patients expected, or wanted blood tests, this was rarely reflected in the patient interviews. For example, doctor 11 perceived that the patient was looking for a decision on further tests, whereas patient 11 was reassured by the normal examination findings and was not expecting tests.

I think that she was coming for reassurance and for a decision as to whether she needs further tests at this stage. (Doctor 11, female, GP partner, 20+ years' experience)	*It didn't enter our head about a blood test… I was glad really that she done the…examined me 'cos I didn't know if there was anything wrong. All I knew was I had the breathing problem. So, I was glad she examined me and she said there was nothing there didn't she? (patient 11, female, 88 years)*

There were no clear examples of the doctor involving the patient in a shared decision to test in any of the interviews (**Table 35**). Although some of the patients said that they thought that they had shared the decision, their description of events suggested that the doctor had made the decision and the patient agreed or acquiesced without evidence of the patient being involved in decision-making. None of the patients recalled being offered any options or alternatives to testing, such as an alternative test or the option of no testing, which is generally accepted to be one of the prerequisites for shared decision-making.

> *I think it was shared. When she suggested it, I said yeah, it's obviously worth checking (patient 6, female, 54 years)*

Patients generally did not perceive blood tests to be a decision where options or choices were possible and there was a lack of demand from patients for shared decision-making.

> *I think you can always say, 'no, I don't want it done', but then if you've got that condition, I mean what choice do you have really? (patient 4, female, 79 years)*

> *Well, you've got to have it done because you've got to find out what the problem is. (Patient 24, male, 79 years)*

In three atypical cases, patients reported that they had asked their doctor directly for blood tests; all these cases had previous abnormal blood test results which they wanted to recheck or monitor, which perhaps empowered or enabled them to make this request. One of the patients reflected that he felt frustrated that he was not able to directly book a test and that he had to wait to get the doctor's 'permission' first, reflecting an imbalance in power and control over decisions around blood tests.

> *I mean I don't see any reason why I can't phone the receptionist and say I want a blood test and she'll say well I can't book you in because you haven't had permission from the doctor, whereas if someone in my case and someone who regularly is going to have blood tests, it should be notified on the computer and the receptionist can book you in and speed it all up wouldn't it? (Patient 24, male, 79 years)*

Table 35: *Decision-making*

Code	Definition	n=28	Example quotes
Doctor led decision	Doctor/patient agree that the decision to test was led by the doctor	24	*I went to the doctor and then she decided to ask me to go for a blood test (patient 25, male, 59 years)* *I get told to go to them, so I go to them (patient 3, female, 20 years)* *I would say that with the blood test was initiated on my decision really, rather than hers (doctor 11, female, GP partner, 20+ years' experience)*
Patient led decision	Where testing was instigated by patient request	3	*I request to them this blood test because when I had my second child, I had inflammation, so I just want to check- (patient 22, female, 44 years)* *Then I asked if I could have a blood test. (patient 24, male, 79 years)*
Shared decision-making	Shared decision – where doctor and patient felt that a joint decision was made. If the doctor makes a decision and the patient agrees this is not shared. Need some evidence that doctor involved the patient in decision-making	0	
Unclear	No information about decision-making, or doctor and patient interview conflicting.	1	*So, my GP said, Dr M, she recommended it. (patient 13, female, 22 years)* *So, her background seemed to be that she was concerned that she might have an underactive thyroid… Yeah, I think [it was a] shared decision. (doctor 13, female, GP partner, unclear/ conflicting with patient 13)*

5.5. Information sharing

Whilst decision-making was a single step, shared understanding required a process of information sharing before, during and after testing.

5.5.1. Sharing information on reasons for testing

Overall, there was a lack of understanding amongst patients about which tests they were having done and why, with 17 out of 28 patients saying they either did not know or were unsure why tests were being done (**Table 36**).

> *You walk out and you see somebody, oh what have, you know, you'll see a friend and they'll say what have you been in there for? Had to have blood tests. What for? Don't really know (laughs). And you don't (Patient 9, female, 72 years)*

Although the lack of shared decision-making was generally perceived as acceptable, the lack of information sharing was perceived as less acceptable by patients.

> *I like to be told 'ok, we'll run some blood tests, we'll have a look for this, that and everything else', but again unless I suppose I'd said to him what are you testing for, I suppose the responsibility, the onus is on me. I sometimes think, like I said, it's like the Secret Service… because the doctors tend to, I don't know. How can I put it? It's almost like they kind of shut you out, you- Oh we'll just check your levels. (Patient 9, female, 72 years)*

> *All I know is I'm just given my four vials of blood and I'm just like where's its going and what's it for, I don't know what you're looking for (Patient 21, female, 29 years)*

Patients saw phlebotomists as another source of information about blood tests. Although phlebotomists could generally explain which tests were being done, my data does not indicate they were able to explain why.

> *I just asked (the phlebotomist) what the blood tests were looking for and how they would show…but they weren't a hundred percent sure what they were. They said what the blood tests would test for, but they didn't really say what they were looking for, if that makes sense. (Patient 27, male, 19 years)*

Table 36: *Information sharing*

Category	Definition	n=28	Example quotes
1. Does patient know why tested?			
Yes	Based on patients own perception of whether they know the reason for testing. Patient has some understanding of what the tests are looking for e.g., aiding diagnosis, guiding treatment, excluding disease, 'baseline'.	11	*This is for the giant cell arteritis because apparently you get inflammation, and it can cause blindness. (Patient 4, female, 79 years)*
No	If the patient says they don't know what the tests are looking for or why they were done.	10	*I had a phone call to say would I have a repeat [test] a month later but I wasn't told why, but I didn't ask why so that's probably my fault. (Patient 1, female, 83 years)*
Unsure	If the patient says they are uncertain, if they offer a guess or use tacit knowledge to say why they think the tests were done (rather than reporting what the GP told them)	7	*So, I don't really know [why] and I suppose Dr M, I said what do you think and he said well we'll do a blood test to confirm and to confirm I presume he meant whether I had this poly fibromyalgia. (Patient 2, female, 82 years)*
2. Does patient know which tests were done? *			
No	Patient only uses vague terms to describe tests e.g., 'bloods', 'blood tests', but cannot give any more detail.	10	*Interviewer - And do you know which blood tests are checked?* *P I'm assuming practically everything (Patient 1, female, 83 years)*
Category	Patient knows the category of tests e.g., 'liver function', 'kidney function', 'markers of inflammation'. Include where they just describe testing organs e.g., 'your kidneys', 'thyroid'.	12	*It was just for renal to check in your… sort of renal, waterworks, you know (Patient 20, male, 76 years)*
Specific named test	Use of specific named tests e.g., 'CRP', 'PSA'.	7	*For bloods they're checking to see if I've got an above average white blood cell count, just to see if there's an infection or something. (Patient 15, male, 18 years)*
Diagnosis rather than test	Patient names diagnosis to describe test e.g., 'check for anaemia', 'coeliac test', 'blood test for diabetes'.	1	*A rheumatoid arthritis test, I think. (Patient 19, female, 52 years)*

Unsure	Unsure if the patient says they are uncertain, if they offer a guess or use tacit knowledge to say which tests were done (rather than reporting what the GP told them)	1	*And did you know which blood tests you were having done?* *P No. I didn't.* *I Did you know anything about what was being detected in the blood test?* *P No. Well, I assume that it would be white blood count and possibly something similar (Patient 28, male, 49 years)*
3. Does patient know anything about limitations of testing?			
No	No awareness or understanding of the limitations of tests	27	*Interviewer - And have any of the doctors talked about any benefits or limitations of blood tests to you?* *Patient - Only to- No, I only think just to find out what is wrong, that's all. (Patient 4, female, 79 years)*
Yes	Some understanding of the limitations of tests e.g., false positives, false negatives. Mentioning that the diagnosis may still remain unclear after test does not count as explanation, unless there is a specific mention of this in context of the test e.g., "this test may not rule X out".	1	*It could show up levels of inflammation, which is not a positive thing for that diagnosis, but it could mean that there was something like that. (Patient 6, female, 54 years)*

**Total adds up to >28 because three patients were aware of some categories of tests and also some specific named tests being done.*

GPs reported sharing more information with patients than was reflected in the patient narratives. For example, doctor 11 expected her explanation of the tests to give the patient a much better level of understanding than the patient reported.

I think she would have understood when I said some people are breathless because of anaemia and I said I'm checking your full blood count and your iron. I think she would have known that that was looking for signs of potential anaemia. (doctor 11, female, GP partner, 20+ years' experience)	*Interviewer Do you know what they're looking for in those tests or what they might pick up?* *Patient All I understood was the inflammation in the blood.* *Interviewer And what do you know about- do you know what that means or-?* *Patient No I don't know, no.* *(patient 11, female, 88 years)*

Although patients commonly mentioned that they were having tests for 'inflammation', few had a clear understanding of what was meant by this. Patients had very little awareness of potential limitations of testing, indeed even the concept of tests having pitfalls or limitations was one that they were mostly unaware of.

> *Interviewer* *Have any of the doctors talked about any benefits or limitations of blood tests to you?*
>
> *Patient* *Only to- No, I only think just to find out what is wrong, that's all. (Patient 4, female, 79 years)*

In contrast when discussing decision-making around blood testing GPs commonly discussed weighing up and considering limitations of testing.

> *As I say, you know, it is that can of worms isn't it? When you do a test and you think why the hell did I put a tick in that box and then it just start- and then almost an avalanche of things, so a raised PV then could end up with BJPs, calcium, serum protein electrophoresis, so more and more tests that I'm less and less experienced with. (Doctor 2, male, GP partner, 20+ years' experience)*

5.5.2. **Sharing test results**

Most patients were told whether their tests were normal or abnormal, but few knew the actual results (**Table 37**). This reflects a paternalistic position, as highlighted by one of the patients.

> *I mean you say you've got your results back; I haven't got my results back, I've got a doctor's interpretation of my results. (Patient 23, female, 24 years)*

Some patients felt that it would be helpful to have copies of the written or printed test results as actual numbers, but this needed suitable explanation.

> *It would be nice if you had the results of the blood test written down but in a way that you could understand what it meant. Like if they said like blood sugar should be under 42 or something like that, but yours is this so do something about it. (Patient 4, female, 79 years)*

Patients generally wanted more information about the meaning of their test results, in terms of what was causing their problems and what they could do about this.

> *What I want is direct input. I want the doctor to tell me ok this is what it means and so this is what I suggest you do (Patient 25, male, 59 years)*

All patients who had abnormal test results (n=7) had some explanation and understanding of the meaning of their results. By comparison, only 3 out of 11 patients with normal results had any understanding of the meaning of the results.

GPs generally assumed that normal test results needed less communication and explanation, so patients were either texted with the results or GPs would leave it to the patient to make contact.

> *I've kind of left the ball in her court now whereas before I think if they were raised and shown signs of inflammation then I would probably make an active plan for follow-up. (Doctor 18, male, GP partner, 0-5 years' experience)*

Four out of seven patients whom the GP reported had 'borderline' test results were not aware of these borderline abnormalities and were under the impression that their tests were normal.

Her plasma viscosity was 1.87 so a little bit raised, there or thereabouts. (Doctor 2, male, GP partner, 20+ years' experience)	*Well, he just said the blood tests came back and it was perfectly normal. No problem, he said no worries at all. (Patient 2, female, 82 years)*

Two patients did not receive their test result at all; both were having routine monitoring bloods for chronic conditions and said that they would not usually contact the practice for results, so were not followed up further. One patient was uncontactable for follow up.

Table 37: *Sharing test results*

Code	Definition	n	Example quotes
Patient awareness of test result (n=27*)			
Results only	Numerical value only	0	
Assessment only	Assessment of result e.g., normal / abnormal	24	*So, I didn't actually get any specific numbers and she- so with the Vitamin D result she said oh no actually your levels are really, really low, I am going to have to prescribe you some- But with the CRP she just says yeah, you've got a raised CRP. (Patient 13, female, 22 years)*
Result and assessment	Numerical value and assessment of whether this is normal/abnormal	1	*Apparently my blood levels… not my blood levels, oh whatever it is, it's going up, it's improving all the time. It's gone from 9.5 to 9.8. (Patient 24, male, 79 years)*
No result	Patient unaware of test result	2	*So as things are, I haven't heard anything back on the back of having them taken at all. So, I presume nothing alarming there but likewise I also don't know whether there is anything positive there either. (laughs) (Patient 8, male, 37 years)*
Patient understanding of meaning of result (n=25)**			
No understanding	Patient is aware of result +/- assessment e.g., 'low', 'high', but no understanding of what that means e.g. pathophysiology, aetiology, diagnosis	12	*Well obviously it's good to know that nothing hideous has been brought up in any of the blood tests but I have absolutely no idea what the blood test was for in the first place so it's kind of a bit like ok, so they're just now saying it's just nothing's come back or nothing's come up sort of thing, saying that its normal but I'm like what's normal? I didn't know what the blood was for, I don't know what the results really tell us (Patient 21, female, 29 years)*
Some understanding	Patient has some understanding of what the test results mean for the patient, beyond assessment	13	*They just explained that everything like came up to no concerns and that, but that they explained about the protein levels being a bit high because of the cold I've just had*. (Patient 16, male, 19 years) *So they explained that from the results they could see signs of an infection so then from there they recommended that I go on a course of antibiotics (Patient 15, male, 18 years)*

**n=27, one patient was uncontactable for follow up, so not included*

***n=25, one patient uncontactable, two patients did not receive their test results, so not included*

5.6. **Consequences of testing**

Tests had medical consequences such as further investigations, treatments, or referral, but also important non-medical consequences which could include anxiety, reassurance, uncertainty, or frustration. Testing allowed doctors and patients to construct meaning and explanation of patients' symptoms, but this required a shared understanding of testing. Without meaning and explanation and without a shared understanding, tests could lead to increased uncertainty and anxiety. The consequences of testing therefore depended on doctor-patient expectations and communication. This is illustrated by paired quotes from case 4 and 26; in both cases the GP felt that the normal result provided reassurance or 'progress' for the patient and doctor, whereas the patient felt anxious that the cause of the symptoms had not been found.

So yeah, I think they [test results] were reassuring for the patient and they were reassuring for me and the rheumatologist that we weren't missing anything potentially serious. (Doctor 4, male, GP partner, 10-20 years' experience)	*Well I'm, you know, I get worried really as to why I'm getting these things and nobody knows why. Obviously the blood tests can't be showing anything. I mean I've had MRI scan a while back, but nothing came of that really. I'm just a bit worried as to why I'm getting these things. (Patient 4, female, 79 years)*
So his CRP had come back normal, a coeliac screen and his faecal calprotectin as well was also normal, so then we're sort of able to progress in thinking this is most likely IBS that's causing this type of pain and talk about ways to manage that. (Doctor 26, female, locum, 0-5 years' experience)	*It's reassuring obviously that it's told me the result that it's not something but it's not reassuring in a way because I don't know what it is and neither does my GP. (Patient 26, male, 25 years)*

Where differing beliefs and expectations were not resolved through communication, this could lead to a distorted understanding between doctor and patient. For example, one patient, having blood tests for joint symptoms and possible arthritis, assumed that negative blood test tests had ruled out cancer.

> *They said it all seems to be clear, but I was worried about cancers, but obviously there's no cancers going on otherwise they would have told me if there was. You would have thought, so wouldn't you? (Patient 25, male, 59 years)*

A lack of shared understanding was particularly problematic for symptomatic patients with normal test results who were left without explanation for their symptoms.

> *I definitely don't feel that a normal result is a win for me by any means and I can't continue to feel the way that I do so I'm just going to have to keep on looking into it and getting that sorted and just kind of being as honest about my symptoms and things with the doctors as I can so that they know where to go with it, I guess. (Patient 21, female, 29 years)*

The fact that the doctor had recommended testing could increase patients' expectations that there must be an underlying problem causing their symptoms, normal test results could therefore, paradoxically, lead to increased expectation that further tests are needed.

> *I mean obviously the blood test was saying that my levels are better so… I suppose there should be other tests then 'cos they haven't found out what's causing the pain and the discomfort and everything. (patient 7, female, 60 years)*

Some patients also felt a sense of frustration, not only that test results did not provide answers or explanation for their symptoms, but also that normal test results in the context of ongoing symptoms seemed to invalidate their experiences, making them feel as if they were being dismissed or written off.

> *Its semi-frustrating because you think well that's another thing that doesn't give me the answer then, that's another reason for them to go there's nothing wrong with you (patient 21, female, 29 years)*

Whilst doctors generally perceived that normal test results need less explanation than abnormal results, these findings suggest that the need to construct meaning and shared understanding is just as important with normal blood test results.

Doctors generally assumed that normal results were reassuring, whereas patients were often actually reassured by abnormal results, even when the abnormalities which were picked up were unlikely to be related to the presenting problem.

> *It was like oh I have a surprise deficiency. (laughs) Obviously I'm really pleased that that was found (Patient 23, female, 24 years)*

5.7. Barriers to shared understanding of blood testing

There was a perception from both doctors and patients that sharing information about testing was the right thing to do '*in an ideal world*', but multiple barriers to information sharing were identified by both doctors and patients. These could be categorised into barriers around knowledge, attitudes, and systems of testing.

5.7.1. Barriers to shared understanding: knowledge

Challenges sharing technical and complex information

Both doctors and patients perceived that tests are sometimes too complex or too technical for patients to understand. There was a lack of shared language which meant that even when doctors tried to share information, patients did not always understand what was being said.

> *I can't remember now whether it was full count or this or that, whatever, I don't know what she was taking about to be honest. (Patient 20, male, 76 years)*

One example was the word 'inflammation', which was often used by doctors and patients, despite the fact that many patients did not understand what this meant. Another example was the fact that if written test results were provided, they were generally printed in a format designed for healthcare professionals not for patients.

> *You could have a copy of your results but it don't mean a thing to ordinary people. (Patient 4, female, 79 years)*

> *If they're bad they will explain to me and show me the results in strange numbers and letters…Yeah, I get a print of it with all the- I don't normally know too much of what it means, just letters and numbers but if something's odd then they'll show me what is odd. (Patient 3, female, 20 years)*

Doctors and patients both perceived that too much information could be confusing and unhelpful for patients.

I don't think she probably fully understood the ins and outs of each blood test, I don't think I went through every test detailing exactly what it might show or what it's for 'cos I think sometimes that's probably a bit too much information. (Doctor 9, male, 0-5 years' experience)

I think sometimes you can confuse yourself if you've got too much information you can sort of read too much into it. (Patient 6, female, 54 years)

Testing was often only one small part of the overall consultation and GPs needed to balance out the most important information for the patient to take away.

Sometimes you don't want to bamboozle people with too much and if you want them just to remember one specific thing about the consultation. (Doctor 21, female, salaried GP, 10-20 years' experience)

Some patients reported that they had been given information about their tests, but that they struggled to retain this information. This is in keeping with the observation that the GPs account of information sharing did not always match with patient recollection. The use of unfamiliar technical terms made it harder for patients to retain information.

She's requested a Vitamin D test as well as just the test to… I don't know, she said some initials, but I just forgot what they were. (Patient 17, female, 18 years)

Some patients perceived that it was their responsibility to ask questions if they wanted to know about their blood tests, rather than the doctor's responsibility to give this information.

I like to be told ok we'll run some blood tests, we'll have a look for this, that and everything else but again, unless I suppose I'd said to him what are you testing for, I suppose the responsibility, the onus is on me. (Patient 9, female, 72 years)

I wasn't told why, but I didn't ask why so that's probably my fault (Patient 1, female, 83 years)

Patients did not always feel able to ask these questions or might not think of the questions until after the consultation.

I definitely would have preferred if they had told me what they were testing for. Perhaps if I had actually asked maybe then they would have but I didn't know I needed to ask until I realised that I didn't. (Patient 14, female, 21 years)

Lack of resources for information sharing

Although doctors and patients discussed using resources such as leaflets and websites to share information about diseases, resources to share information about tests were not used in any of the cases.

> *I use a lot of patient.co.uk handouts if I know they've got a condition or I think they've got a condition to give some sort of simple patient information leaflets but I don't usually give them one about blood tests. (Doctor 17, female, GP partner, 5-10 years' experience)*

Patients and doctors both said that they thought that written information would be useful, but doctors were concerned that leaflets about tests could potentially cause anxiety.

> *I think actually, you know, now you say it, I think if we were to have patient information leaflets on the meaning of a plasma viscosity and a CRP that might be really helpful. Having said that, if you have any information leaflet that lists all the things it could be, that leads to anxiety itself in patients. (Doctor 5, female, GP partner, 10-20 years' experience)*

Where doctors discussed information resources for blood tests these were mostly felt to be designed for patients who had received test results, but were often not suitable for shared decision-making around whether to test.

> *I mean labtests online is quite good isn't it but it's not very- It doesn't really go into the pros and cons of having the tests done, it's further down the line isn't it, it's you've had the test done, this is what it potentially means. (Doctor 4, male, GP partner, 10-20 years' experience)*

Some patients said they usually looked up health information on the internet. However, without information about which blood tests had been requested, they were not empowered to look up information on their tests.

> *I know NHS has a website and they have some information on there as well. But I guess I don't know what I'm searching for because I don't know what blood test was being run so I wouldn't know where to start I guess. (Patient 12, female, 27 years)*

Others actively chose not to look up health information online because of the difficulty accessing trustworthy information, and risks of causing anxiety.

I try not to Google because as you know with the internet everything tells you you're dying (laughs), and I didn't want to hear that. (Patient 21, female, 29 years)

5.7.2. Barriers to shared understanding: attitudes

Blood tests perceived as low priority

Doctors perceived that tests were relatively trivial interventions and were therefore low priority for information sharing. Doctors described having multiple other issues to discuss in a consultation and had a perception that patients generally did not want to know much about blood tests.

I don't find patients ask terribly much about what the blood tests are that we're testing for… I don't have a massive feeling that they want to know terribly much more but maybe I'm wrong about that, I don't know. (Doctor 15, female, GP partner, 10-20 years' experience)

Paternalism

Some doctors and patients felt that testing is an area which justified a more paternalistic approach. The imbalance of knowledge and power between doctors and patients meant that patients did not think it was possible for them to engage with decisions around testing. Patients reported 'trusting' their doctors to make the right decisions on their behalf, and this was seen as a reason not to ask further questions.

I kind of trust them enough to know I know they know what they're doing therefore I don't really like need to have too much involvement in what they're- not necessarily what they're looking for but what they're doing, I trust them. (Patient 3, female, 20 years)

I think that's still one of those areas of being a GP that patients just trust you're going to ask for the right tests. (Doctor 15, female, GP partner, 10-20 years' experience)

The perception that testing is '*still one of those areas*' seems to reflect an awareness amongst GPs of the changes over time in medical decision-making, and the

gradual move away from paternalism. Some doctors reflected on this during the interviews and felt that this was an area where they could try to improve their practice.

> *I'm probably not as good as I might be at sharing that decision, so it's probably more mine actually but I think I will reflect on that. Yes, I'm probably, you know, I think our GP trainees and our trainers are very good at doing shared decision-making. I'm probably perhaps a bit more old-fashioned in my approach. (Doctor 6, female, GP partner, 5-10 years' experience)*

Justification for paternalistic approach: protecting patients from anxiety?

Doctors perceived protecting patients from undue anxiety was sometimes a justification for withholding information about blood tests. Doctors perceived that minor abnormalities which were not clinically relevant, could cause anxiety if shared with patients.

> *The difficulty with that is that there are lots of minor variations… you see all the red exclamation marks that come up when you get your blood results back and whilst we know that the vast majority of those are nothing to be concerned about and we would file that as a normal result or satisfactory, it could cause a lot of concern for patients potentially. (Doctor 8, male, GP partner, 0-5 years' experience)*

Doctors also perceived that sharing possible diagnoses, before test results were available, might cause unnecessary anxiety. Rather than giving details of which tests were being done, or what was being tested for, they might therefore use more general terms.

> *[I said]…there is one condition that can give you both pain with that, so I want to do some extra tests to just rule that out. And that's thinking about myeloma. I don't think I said to her the word myeloma or what it was necessarily. (Doctor 7, female, GP partner, 5-10 years' experience)*

In this case the possible diagnosis being checked for, myeloma, is a type of cancer, but the doctor avoided mentioning this explicitly. Rather than

mentioning the word 'cancer', doctors and patients tended to allude to possibilities of something 'serious' or 'worrying'.

> *I probably sort of don't say exactly what I'm looking for, maybe saying make sure there's nothing else worrying going on is probably what I'd be saying more. I think a lot of people tend to understand what that means. I mean a lot of people are worried about cancers and things like that. (Doctor 24, female, salaried GP, 0-5 years' experience)*

However, even when these possible serious diagnoses were unspoken, or only alluded to in the consultation, patients used guesswork or tacit knowledge to infer the possibilities.

> *Interviewer* What do you think they might be looking for in those tests?
>
> *Patient* Something in the blood, like this time I thought maybe because I've got lumps in my neck, I thought maybe leukaemia 'cos I know that's something, but again that's just a guess.
>
> *Interviewer* Was there anything like that mentioned?
>
> *Patient* No. It was just to look out for symptoms like night sweats, weight loss and just feeling unwell.
>
> *(Patient 12, female, 27 years)*

Some patients also perceived that doctors 'couldn't' share information with them until they had definite answers.

> *I imagine my GP is on the ball, I've always thought she was anyway, but as I say obviously there's more to this than meets the eye, more to this than I've been told. They can't tell me anything till they know for certain, can they? (Patient 1, female, 83 years)*

Some patients felt that this was justifiable, and would rather doctors did not share possibilities or uncertainties as they wanted to avoid unnecessary anxiety.

> *I'd rather not know possibilities; I only really want to know definites… because there's no point worrying yourself sick about something that you might not have but you can't stop yourself. (Patient 23, female, 24 years)*

Other patients wanted more information about the possible diseases or diagnoses before testing, so that they could be aware not just what diagnoses had been made, but also about what possibilities had been ruled out.

> *I think it's just good to know what you're looking at… It would just be good to know because then once it comes back negative you can chuck that away, it's gone, its eliminated but if there is a tick list of things that it could be, it's good to be able to cross those things off …So the more information they can give people about what they're looking into the better I think personally. (Patient 21, female, 29 years)*

Another justification for holding back information about testing, and taking a more paternalistic approach, was the perception that tests were sometimes done for doctors' purposes, '*sort of behind closed doors*', that they were not meant for patients, and that doctors needed to have thinking space to work out what was going on before they could share this with patients.

> *No, I mean it's difficult to say but plasma viscosity is one of those ones that I often use to reassure myself to look for things but honestly, sometimes it's one of the ones that I try not to discuss in too much detail with patients because I think generally it can lead to more alarm. It sounds awful but the interpretation of it is so difficult, I perhaps downplay it until I know it's a significant thing. (Doctor 5, female, GP partner, 10-20 years' experience)*

> *In my head the results have always been for the doctors, they're not for me, do you know what I mean? 'Cos, they tell you whether you have anything to worry about or anything to follow up with but you're never going to understand what those things are so really it seems to be for the doctor's eyes only. (Patient 21, female, 29 years)*

However, although information about testing was sometimes withheld by doctors to prevent patient anxiety, some patients perceived that a lack of information sharing was more likely to provoke anxiety.

> *I worry more by not knowing. I do personally. I prefer to know. I think right, I'm sort of like- then I'm prepared, aren't I? I can sort of get all my ducks in a row (laughs). (Patient 9, female, 72 years)*

5.7.3. **Barriers to shared understanding: systems of testing**

Another potential barrier to shared understanding was the systems of testing, particularly around how test results were communicated.

Uncertainty and variation in test result communication

Patients could receive their results face to face, by telephone, text message or by letter, and communication could come from a doctor, allied health professional or from receptionists. This communication of test result could be instigated by the GP practice, or it could be up to the patient to initiate communication. Most doctors made individualised decisions about how to share results depending on their knowledge of the patient, the clinical context, and the test results. Methods of communicating results varied between doctors and were based on habits, unwritten heuristics, and personal preferences rather than protocols.

> *There aren't protocols that we use. There's a lot of debate in the practice as to how we manage blood test results in that the onus is always put on the patients to call about the results of the blood tests. So, it depends what the results show. If there's any significant abnormality that needs action, we normally speak to the patient straightaway. (Doctor 8, male, GP partner, 0-5 years' experience)*

> *It's always difficult when people in a practice do different things or if I'm away, so somebody else is filing my results they may do it in a different way. Yeah, I think it's up to the individual. (Doctor 17, female, GP partner, 10-20 years' experience)*

As a result of this variation patients often struggled to navigate the system and often used guesswork to decide whether or when to contact the practice for results.

> *Certain tests can be given to you via text message. If there is, well sometimes its sent to you via paper, letter form, and if there's any cause for concern sometimes the reception will call you and say can we book you in with, it really does, it differs so much as a change in different methods really for every different type of test, so I just kind of go oh ok, I haven't heard for a certain amount of time, I'll call up the reception. (Patient 21, female, 29 years)*

Although doctors often assumed that patients would contact the surgery for their test results, this was not always the case. For example, in case 8 the doctor had recorded in the notes that the patient should make an appointment to see a doctor, relying on the patient to contact the practice to receive this message. The patient, who has regular blood tests for monitoring, admits he has never phoned

for his test results. Instead, the patient assumed that the practice would contact him if there was any problem.

His CRP is 114, she's put on there see doctor if still has the symptoms. I think I probably would have contacted him with that CRP result …But I'm sure, I know he would call if he was deteriorating anyway (Doctor 8, male, GP partner, 0-5 years' experience)	*I've never, ever asked for my test results. I've always just turned up, had my blood taken, gone away and always with the assumption that if there was anything wrong someone would let me know. (laughs*) (Patient 8, male, 37 years)

Similarly, in case 20 the doctor put a comment to repeat the test, assuming the patient would contact the practice to receive this message. When the patient did contact the practice, he felt aggrieved that he had not been contacted.

I didn't speak to him, I just put a comment that they were all improving and to repeat in two weeks' time. (Doctor 20, male, GP partner, 0-5 years' experience)	*So anyway, I waited a week, went there last Friday and asked the lady and she said yeah it was all clear, oh but they wanted you to have a blood test again on one particular thing and I'm thinking well I would have never known that if I hadn't had come and asked…. How many people ring up and ask for their results, how many people make the effort to go down and ask, it's a bit…Yeah, I thought well it's a bit lax. (Patient 20, male, 76 years)*

Doctors generally expected patients to know how to access their test results, but patients were often unsure about the best way of doing this.

I always say you must get a result one way or another hearing that it's either normal or abnormal, you need to make sure you contact us for a result if it hasn't come through to you. (Doctor 13, female, GP partner, 10-20 years' experience)	*I don't know how I do that actually. Maybe I ring up and- probably ring up and just ask the receptionist how I'd go about doing that, 'cos I don't know if I need like a whole appointment for that, I don't know if they could send those [blood test results] to me. I don't know how it works. (Patient 13, female, 22 years)*

Communication of results by non-clinical staff

Another barrier to shared understanding was the system of communicating test results via receptionists or non-clinical staff, who were unable to give explanation of the meaning of test results.

The receptionist said, I told her why I was phoning, I said I haven't had the results of my blood test and she said yes, oh yes, I've looked it up, everything's fine, no further action. So, I didn't need to go back to see the doctor and then you think well in a way you've got a closure of a sort but not of what you'd originally perhaps come about… So, I don't know. (Patient 18, female, 76 years)

Some patients felt unhappy about receiving test results from non-clinical staff members who might not have the training to understand test results.

Chasing results is difficult to get from the secretary. No disrespect to them either but that's a very difficult situation 'cos she's having to discuss what she sees on the screen. And are they right to do that and are they right- do they understand that? I don't know. (Patient 5, male, 43 years)

I don't even bother- I could ring here and find out the results but I don't know about the receptionists and I'm sure they're well trained, but do they actually know how to read the blood results? (Patient 10, female, 56 years)

Text messages

Whilst systems of texting patients about their test results was generally perceived positively by doctors, there were mixed views from patients. Some welcomed this as a quick and easy way to get reassurance with normal test results.

That's suits me 'cos I know… the doctor explained that if they were normal results, they'd come through text so then I knew when I got the text and I quickly read through I was like oh its fine. But obviously if it was abnormal, I wouldn't want to receive that by text. (Patient 12, female, 27 years)

Others felt that text messages did not really convey sufficient information or explanation to allow an understanding of the meaning of these results.

I think having a text saying everything was normal sort of reassured me that nothing greater was going on, but also again having it saying it was normal was a little bit like well I'd kind of like a little bit more explanation. (Patient 27, male, 19 years)

Text messaging systems were generally designed to prevent two-way communication, making them efficient for doctors. However, when doctors included safety-netting advice asking the patient to get back in touch if they had

concerns or questions, patients perceived that they had no clear route to communicate back to the doctor.

> *I find it quite new now to be texted by my doctor and its very much I feel a one-way method of communication because I can't reply, and I can't just text my doctor back. So, my doctor has said several times in texts if this hasn't worked let me know if you have any more questions, please get in contact with me, but the only way I know of doing that is by booking an appointment or a phone call. I can't just post a message on a message board or send a text to that doctor. So that's a confusing communication method for me. (Patient 23, female, 24 years)*

Some patients even felt that a text was perhaps an inappropriate, or 'flippant' way of communicating test results, particularly for those with more complex, ongoing problems.

> *I'd say in people's situations like mine where it's an ongoing thing, I don't feel that a text message is sufficient… I feel if there's more investigations to be done from that point on, just sending a blunt text isn't really sufficient because it means nothing to me… So, I don't know, it just feels a bit, sometimes when you get a text about something like that it seems a tad flippant. (Patient 21, female 29 years)*

This contrasted with doctors' perceptions of patients, most of whom assumed that patients were very happy to receive text message communication.

> *I love it. I think it doesn't work for all patients and it's good that there's an opt in/opt out option but I think particularly our patients they're very proactive, they like to receive communication from us and they quite often are happy to be texted. (Doctor 12, female, salaried GP, 0-5 years' experience)*

> *The patients love them, generally the feedback's been very good. (Doctor 27, female, locum, 0-5 years' experience)*

Multiple issues relating to primary care systems which are not specific to testing were also mentioned as a barrier to test result communication, including time pressures, access to GP appointments, lack of continuity of care, and GP workload. These issues have been widely discussed in the academic literature so have not been analysed further.

5.8. What do patients want to know about their blood tests?

Although the complexity and technical nature of tests was perceived to be a barrier to information sharing, patients did not necessarily want technical information, but they did want to know what was being tested and why.

> *I mean I know obviously it's probably very technical and it's not my expertise at all, but they could just say this is what we're checking, this is what we're hoping for and x, y and z really. They took I think three vials of blood and I don't know which ones (laughs). (Patient 26, male, 25 years)*

By building a shared understanding of the reasons for testing and a shared expectation of tests, doctors and patients could create meaning from the result. Case six illustrates a successful example of this; the GP shared her expectations of testing with the patient, explaining that the tests might be normal and what that would mean. She then also contacted the patient to discuss the meaning of the normal test results.

> *So I said to her if they're completely normal I'll be very happy that it's not this condition. If they are very raised, I will phone you and we'll treat you with steroids but I expect they may well be normal. So I kind of gave her the expectation that they probably would be normal and I phoned her and what I've written is 'phoned to reassure inflammatory markers normal, she's happy to wait and see how things go' because she was feeling slightly better after resting on her holiday. (Doctor 6, female, GP partner, 5-10 years' experience)*

The patient in this case reflected on the importance of creating meaning through a shared understanding of testing.

> *I think as long as people know what the tests are for, it's not just oh we'll do a blood test. It's just making people aware of what they're testing for and then people can understand… because blood tests don't really mean anything unless you know what it's for does it really? (Patient 6, female, 54 years)*

5.9. **Summary**

The first objective of the qualitative research described in this chapter was to explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests. My findings demonstrate a mismatch between patients' and doctors' expectation and understanding of testing. Patients were frequently unaware which tests have been done and why. Patients had high expectations of tests; expecting them to provide clear-cut answers, whereas doctors' expectations were more modest.

The second objective was to provide in-depth exploration of patients' experiences of testing, from GP consultation to results. I found that none of the patients interviewed engaged in shared decision-making around testing, nor did they perceive blood tests to be a decision where options or choices were possible. Whilst decision-making was a single step, shared understanding required a process of information sharing before, during and after testing. 'Normal' test results required explanation and understanding to generate meaning and reassurance for patients; results without understanding could lead to further uncertainty and anxiety and for some people led to a feeling of being dismissed or written off.

The final objective was to identify barriers and facilitators to improved communication, in order to inform improved communication in future. Barriers to shared understanding were identified in areas of knowledge, attitudes, and systems of testing. Barriers in knowledge included the perception that tests were complex, technical, and difficult to explain, and a lack of resources for information sharing. Barriers in attitude included perceptions that tests were relatively trivial interventions and therefore low priority for information sharing, and that tests were an area where a more paternalistic approach could be justified, particularly when sharing information could potentially cause anxiety. Finally, barriers in systems of testing included uncertainty and variation in methods for communicating test results, and the fact that test results were

frequently communicated by non-clinical staff, or by text message, which could leave patients with unanswered questions and no clear route for addressing this.

In the next chapter I will discuss the strengths and weaknesses of this research, comparisons to existing literature and the clinical implications of these findings, alongside the quantitative research.

CHAPTER 6. DISCUSSION

6.1. Chapter overview

In this chapter I discuss the findings of my quantitative and qualitative research. First, I discuss the overall findings, linking these back to my thesis aims and objectives. Next, I compare these findings to the existing literature, then I discuss the strengths and limitations of my research. After that, I discuss the implications for research and clinical practice, before considering how the qualitative and quantitative research components of the thesis can be synthesised and making some final personal reflections.

6.2. Overall summary of findings

The overall aim of this thesis was to explore the diagnostic utility and clinical practice of inflammatory marker testing in primary care, including how results are shared with patients, using a mixed methods approach. I will summarise the key findings of my research by referring to the thesis objectives as outlined in the background chapter **1.9**.

6.2.1. Summary of quantitative results

For the quantitative component of the thesis there were five main objectives. Results are summarised under each of these objectives.

Objective 1: Epidemiology of inflammatory marker testing

Objective 1 of my thesis was to describe the baseline characteristics in terms of age, gender, socioeconomic status, and ethnicity of patients having inflammatory marker tests in primary care.

Inflammatory marker testing was more common in females, in white ethnic groups, and in the more affluent. Conversely, abnormal results were more common in patients from more deprived socioeconomic groups. This is in keeping with the inverse care law,(209) with relative over-testing in the affluent, and relative under-testing in more deprived groups. Higher testing rates may also in part reflect higher consultation rates in certain socio-demographic groups; for example, women represent 60.2% of GP attendances, and comprise 61.8% of the inflammatory marker tested cohort.

Objective 2: Inflammatory markers and overall disease outcomes

Objective 2 was to determine the diagnostic accuracy of CRP, ESR and PV, singly and in combination, for relevant disease (infection, autoimmune conditions, and cancers) in primary care, and compare disease incidence in tested versus untested populations.

Overall incidence of relevant disease was 15.0% in the raised inflammatory marker group, in comparison to 6.0% in the normal inflammatory marker group and 3.4% in the untested comparison cohort. In patients with a raised inflammatory marker, the most common diagnoses were infection (6.3%), followed by autoimmune conditions (5.6%) and cancers (3.7%). Disease incidence was higher in those with more than one simultaneous raised inflammatory marker (22.6%) and in those with a second repeat test which was persistently raised (23.8%). Disease incidence increased with rising inflammatory marker levels in a dose-response relationship. The overall sensitivity of inflammatory markers for relevant disease was low (<50%), meaning over half of patients with

relevant disease were missed by the tests. Inflammatory markers are therefore not suitable as a 'rule-out' test, which would require high sensitivity to avoid missed pathology.

To explore comparative accuracy of the three inflammatory marker tests, area under receiver operating curves (AUC) were compared for any relevant disease for those with two tests performed simultaneously. Small differences were seen with AUC ranging from 0.659–0.682. CRP had the highest overall AUC, largely because of marginally superior performance in infection (AUC CRP 0.617 versus ESR 0.589, $p<0.001$).

Overall, 39.4% of the tested cohort had more than one inflammatory marker tested simultaneously; mostly CRP and ESR (51,949), followed by CRP and PV (11,107). Testing multiple inflammatory markers simultaneously was associated with more abnormal and discordant results; when a single inflammatory marker was tested, 25.9% were found to be raised, when multiple inflammatory markers were tested, 36.3% had one or more raised inflammatory marker (14.4% concordant raised values and 22.0% discordant). Adding a second test led to a tiny increase in discriminatory ability for relevant disease measured by the AUC (CRP 0.682 versus CRP+ESR 0.688); whilst the p-value reached thresholds generally considered to be statistically significant ($p<0.001$), this was very unlikely to be clinically significant. The negative predictive value (NPV) of a single inflammatory marker (94.0%) was effectively the same as the NPV of combined inflammatory markers (94.1%), suggesting that testing multiple simultaneous inflammatory markers does not help to rule-out serious pathology. No combination of inflammatory marker tests can be used to rule-in or rule-out disease confidently, with a maximum sensitivity of 60.6% for the combined test defined as positive if either CRP or PV were raised.

Objective 3: Symptoms and cascade testing

Objective 3 was to determine the symptomatology of patients with inflammatory marker testing in primary care and measure the consequences of testing in terms of numbers of consultations, blood tests and referrals.

The most common symptoms associated with inflammatory marker testing were abdominal pain, tiredness symptoms, joint symptoms and infective symptoms. Non-specific symptoms such as tiredness, dizziness and low mood were relatively more common in the normal inflammatory marker compared to raised inflammatory marker groups, indicating that these non-specific symptoms are less likely to generate raised inflammatory markers. Raised inflammatory markers were associated with higher rates of GP consultations, blood tests, and referrals compared to normal inflammatory markers. This is likely to relate to the wide range of potential differential diagnoses in patients with raised inflammatory markers. Although the unit cost of inflammatory marker tests is relatively low, the total costs, including these follow-on consultations, investigations and referrals are likely to be substantial.

Objective 4: Inflammatory markers and cancer

Objective 4 was to determine the diagnostic accuracy of inflammatory markers for cancer diagnosis in primary care, including stratification by age, gender, inflammatory marker level and cancer type.

Inflammatory marker tests were not a useful rule-out test for cancer, with a sensitivity of 46.1% for CRP, 43.6% for ESR and 49.7% for PV; however, a raised inflammatory marker could still be a useful clue to cancer diagnosis. Men aged over 50 and women over 60 with a raised inflammatory marker had a one-year cancer incidence of 6.44% and 4.22% respectively - above the 3% threshold considered by NICE to warrant urgent investigation or referral for suspected cancer.(65)

Objective 5: Inflammatory markers and mortality

Objective 5 was to explore the association between inflammatory markers and one-year mortality in primary care.

Inflammatory markers are a strong predictor of all-cause mortality in primary care. The association between raised inflammatory markers and all-cause mortality is seen in all age groups except patients aged less than thirty years. Men with raised inflammatory markers have a higher one-year mortality than women (9.78% vs 5.29%). Of the three tests examined, CRP has the highest predictive accuracy for mortality. A logistic regression model containing age, gender and CRP test result had an AUC of 0.89. Whilst these findings are interesting from a research perspective, the clinical implications of these findings are less clear-cut, as discussed further in section **6.5.1**.

6.2.2. Summary of qualitative results

For the qualitative component of the thesis there were three main objectives, results are summarised below under each of these objectives.

Objective 1: To explore to what extent doctors and patients have a shared understanding of the use of inflammatory marker blood tests

The results of the qualitative study demonstrated a mismatch between patients' and doctors' expectations and understanding of inflammatory marker testing. Patients were frequently unaware of which tests had been done and why. This has important implications for the legal and ethical principles of informed consent. Patients expected that tests would provide diagnostic certainty and had little or no awareness of any limitations of tests. Doctors were aware of the limitations of testing, but rarely shared this with patients. Most patients were

told whether their tests were normal or abnormal, but few knew the actual results. Whilst doctors tended to be reassured by normal results, patients with ongoing symptoms often perceived that normal results were unhelpful.

Objective 2: To provide in-depth exploration of patients' experience of testing – from GP consultation to results

None of the patients interviewed engaged in shared decision-making around testing, nor did they perceive blood tests to be a decision where options or choices were possible. Whilst decision-making was a single step, shared understanding required a process of information sharing before, during and after testing. 'Normal' test results required explanation and understanding to generate meaning and reassurance for patients; results without understanding could lead to further uncertainty and anxiety and for some people led to a feeling of being dismissed or written off.

Objective 3: To identify barriers and facilitators to communication and shared understanding

Multiple barriers to communication and shared understanding of testing were identified. Knowledge barriers included a perception that tests were too complex or technical for patients to understand, a lack of shared language, and a lack of resources for information sharing. Attitudinal barriers included the perception that tests were low priority for information sharing and a perception that a more paternalistic approach could be justified when testing, in order to protect patients from unnecessary anxiety. Finally, multiple barriers relating to systems of testing were identified, with systems often unclear and confusing for patients, and based on habits and routines rather than clear protocols. Doctors and patients each often assumed that the other party would make contact regarding test results,

with implications for patient safety. These barriers to shared understanding have important implications for clinical practice, and future research.

Overall, the CPRD research unmasked and quantified limitations and uncertainties inherent in inflammatory marker testing, which are not unique to these tests, but can be a particular challenge due to the non-specific nature of inflammatory markers. The qualitative study showed that whilst doctors were aware of these limitations and uncertainty in testing, this was rarely, if ever, discussed with patients. This led to a lack of shared understanding and shared decision-making. Although doctors perceived that normal test results were useful for reassurance, this was not reflected in patients' accounts of testing.

6.3. Findings in the context of existing literature

6.3.1. Quantitative study

Epidemiology of inflammatory marker testing

The inverse care law has been described for many decades,(209) yet there is limited evidence exploring the influence of socioeconomic status on rates of blood testing. Evidence of lower rates of testing, and higher rates of abnormality in deprived populations have been demonstrated for HbA1c testing in patients with diabetes,(210) prostate specific antigen testing and breast cancer screening.(211, 212) Previous studies exploring variation in inflammatory marker testing rates between GP practices used statistical analyses to control for differences in socioeconomic status.(13) This is because socioeconomic status is a strong predictor of multimorbidity and ill health, and therefore some observed differences may be clinically warranted, however this approach risks masking evidence of the inverse care law. The findings in this thesis that testing was less frequent and abnormal results were more common in those with higher

deprivation measured by postcode IMD, suggests that higher testing rates do not necessarily match the populations with higher rates of disease. It is unclear from these results to what extent this reflects under-testing in more deprived populations, or possible over-testing in more affluent populations, as previously suggested by Spence.(213) Future research into diagnostic testing and the inverse care law could help to elucidate this further.

Inflammatory markers and overall disease outcomes

Most studies of inflammatory markers, which I have reviewed previously,(1) consider single diseases, and most have been based in secondary care. The evidence in this thesis confirms the association between inflammatory markers and inflammatory conditions; however, the PPVs are lower in a primary care population with its lower prevalence of these conditions.

This research has shown that inflammatory markers have an overall sensitivity of <50%, which means they are not a useful test of exclusion. This is discordant with the practices of GPs revealed by my pre-doctoral qualitative research, in which GPs described how inflammatory markers could be used as a 'fishing' test, and to help rule-out disease:(2)

> *So if we had a test that, a single blood test, that doctors could do which would reassure the patient there was nothing bloody wrong at all, then that would be a very popular test. We'll have the "nothing wrong at all" test for you, sir … You know, all the other tests are, well, you might have this specifically wrong with you or you might have this … But the CRP is probably the closest thing that we've got to a "nothing wrong at all" test.*

The finding that 50% of patients with infection, autoimmune disease or cancer had a negative inflammatory marker is therefore at odds with GPs' perceptions of these tests as a '*nothing wrong at all*' test. This is also discordant with current guidelines for chronic fatigue, irritable bowel disease, and suspected dementia, which recommend inflammatory marker testing in order to 'exclude other diagnoses.'(114-116)

Not only is the sensitivity of inflammatory markers poor, but this research has also shown that patients with a normal inflammatory marker are in fact at an increased risk of relevant disease compared to untested controls, demonstrating the risk of false reassurance from a negative test. This is because the mere fact that an inflammatory marker test has been conducted, irrespective of the actual result, predicts disease.(22) This additional risk is only partly eliminated by a negative test result, leaving the negative test group still at a higher risk than those untested (6% relevant disease in test negatives versus 3.4% in the untested). This is in keeping with previous research in primary care patients with a normal platelet count (<400x109/1) who have a 4.1% one-year cancer risk.(189) Similarly, a normal primary care haemoglobin result has been showed to be associated with colorectal cancer (odds ratio 1.5; p=0.001) and a normal chest x-ray is associated with lung cancer (odds ratio 6.9; p=0.001).(214) Likewise, in secondary care settings, the timing and repetition of testing has been shown to be frequently more predictive of disease outcomes than the actual test result.(215)

Interestingly, the one outcome measure which did not follow this pattern was mortality - patients with a normal inflammatory marker had an overall one-year mortality of 1.41% compared to a one-year mortality of 1.62% in the untested comparison cohort. This suggests that the negative likelihood ratio of a negative inflammatory marker test result is greater than the positive likelihood ratio of the clinicians' decision to test, in other words the additional mortality risk associated with the clinicians' decision to test is eliminated by a negative result.

Comparative test accuracy

In a previous systematic review, limited evidence directly comparing CRP and ESR was found for a small number of specialist disease outcomes in secondary care settings, preventing the authors from making recommendations about the preferred choice of test.(144) The authors concluded that a combined CRP|ESR test (defined as positive if either CRP or ESR were positive) had consistently higher sensitivity and lower specificity than individual CRP and ESR tests, in

keeping with the results in this thesis. These findings are entirely predictable, given that a less stringent criteria for a positive test (either CRP or ESR positive) is equivalent to a lowered threshold, which by definition leads to an increase in sensitivity and a drop in specificity. The large number of people with multiple simultaneous inflammatory marker tests in the research presented in this thesis allowed direct comparisons of the area under ROC curves for inflammatory marker tests singly and in combination. These findings confirm that CRP has slightly higher diagnostic accuracy for infections; however, no significant differences were found for autoimmune conditions and cancers. Furthermore, testing multiple simultaneous inflammatory markers did not lead to clinically significant improvements in diagnostic accuracy as measured by the AUC.

Symptoms and cascade testing

The frequency of non-specific symptoms in this study is in keeping with my pre-doctoral qualitative research findings in which GPs described using inflammatory markers in patients with undifferentiated symptoms to help manage diagnostic uncertainty.(2)

The research presented in this thesis identifies a significant difference in the rate of GP consultations, blood tests, and referrals in those with abnormal tests but no relevant disease ('false positives'), compared to those with normal tests and no relevant disease ('true negatives'). This has previously been described by Deyo *et al* as the 'cascade effect of medical technology'.(25, 216) Sah *et al* described this as 'investigation momentum' - whereby an inconclusive or ambiguous test result leads to increased uncertainty and generates additional diagnostic testing which would not have been deemed necessary if the index test had not been performed.(217) Whilst this cascade effect has been measured following prostate specific antigen testing,(217) and in one small primary care cohort of patients with low pre-test probability of disease in primary care,(27) this is the first time it has been demonstrated following inflammatory marker testing. The findings are in keeping with my previous qualitative research in which GPs described

cascades of further tests following a raised inflammatory marker as a particular problem due to the non-specific nature of these tests:

> Then you think suddenly, well should I be looking further and further and further, but that could mean more and more random investigations until you get the point where you goes, oh, I'll just do a whole body CT scan to see if anything pops up I suppose.' *(2)*

Inflammatory markers and cancer

Previous cohort studies in the general population have explored the association between inflammatory markers and risk of future cancer,(64, 111, 218) including colorectal,(129, 219) lung,(128) ovarian(220) and breast cancer.(130, 221) There are also a small number of primary care case-control studies of specific cancers including bladder and kidney cancers, Hodgkin and non-Hodgkin lymphomas, and myeloma, although for single cancers these showed very low positive predictive values (PPVs) for a raised inflammatory marker.(61-63, 222)

Several older studies have examined cancer risk in patients with significantly raised inflammatory markers in secondary care settings. For example in a cohort of 1004 hospital outpatients with an ESR > 100 mm/h, 17% had malignancy,(149) while in another hospital cohort 16% of those with ESR > 100 mm/h had malignancy.(150) In the CPRD cohort presented in this thesis, a comparable 13.6% of those with ESR > 100 mm/h developed cancer in one year, the slightly lower figure possibly representing the primary care setting.

There have been no previous studies to my knowledge measuring the overall clinical utility of inflammatory markers for cancer diagnosis in primary care. CRP and ESR have been evaluated as a tool for predicting cancer in a highly selected cohort of patients with non-specific symptoms referred to Diagnostic Outpatient Clinics for rapid access to cancer diagnostics in Denmark.(223) In this setting the cancer prevalence was much higher at 19.8%; those with raised CRP had an odds ratio of 1.41 for cancer, after adjustment for age and sex, comparable to our adjusted DOR for raised CRP of 1.79.

The findings in this thesis demonstrate that this association exists, not only in the highly selected group of patients referred with suspected cancer, but also in unselected primary care patients, where the initial triaging and referral decisions must be made. Although positive predictive values were clinically useful, the low sensitivity means that inflammatory markers should not be used as a rule-out test. No evidence was found that CRP, ESR or PV were superior to one another in relation to overall cancer detection, however this may not be true for all types of cancer; for example recent studies have shown that ESR and PV are superior to CRP for myeloma diagnosis.(59)

Inflammatory markers and mortality

The association between CRP and mortality is in keeping with population-based studies examining all-cause mortality(134, 224) and cardiovascular mortality,(125) as well as hospital-based studies of patients with specific diseases including COPD,(225) diabetes,(226) chronic kidney disease,(227) pneumonia,(228) cancer,(112, 229) and more recently covid-19.(230) The research reported in this thesis is the first (to my knowledge) to demonstrate that an association between inflammatory markers and mortality is also seen in a primary care setting and over the shorter term. The finding that men with raised inflammatory markers are at higher mortality risk than women may reflect gender differences in healthcare-seeking behaviour in primary care; men have lower rates of consultation, so might be 'sicker' on average when selected for blood tests.

Multiple risk tools have been developed to identify patients with frailty and some of these have been validated by measuring how well they predict mortality.(231) The aims of these tools include helping predict those at risk of unplanned hospital admissions and to allow targeted interventions to people at an increased mortality risk. The National Institute for Health and Care Excellence multimorbidity guidelines systematically reviewed 41 of these risk tools for predicting mortality; the majority were of low or very low quality and a need for

further research in this area was identified.(231) Current risk tools include variables such as disease status, socio-demographic factors and laboratory test results (e.g., anaemia, raised platelets). However, none in current use include an inflammatory marker test. The most commonly used of these risk tools are the electronic Frailty Index(137) and Qmortality.(136) The former has an AUC for mortality of 0.76; the latter an AUC of 0.85 for women and 0.84 for men. More recent research by Deelen *et al* has used combinations of biomarkers to predict mortality across all age groups; 226 potential biomarkers were selected, but CRP, ESR and PV were not considered.(232) Deelen *et al* generated a model using 14 biomarkers with an AUC of 0.837: however, of the biomarkers considered, only albumin is available in primary care, limiting the clinical usefulness of their findings. CRP by comparison is a low cost and widely available test. The evidence in this thesis shows that CRP alone has an AUC of 0.78, increasing to 0.89 for a model including CRP, age and gender. This suggests that inflammatory markers could be a simple indicator with a comparable or better accuracy than currently used mortality prediction tools. Although these findings are interesting from a research perspective, the clinical utility of these findings is less straightforward, and is discussed further in **6.5.1**.

6.3.2. **Qualitative study**

A recent BMJ review article explored communication of laboratory tests to patients; however, patients' perspectives were not addressed and recommendations were based on clinicians' experiences rather than empirical evidence.(233) The qualitative research presented in this thesis helps build this empirical evidence and identifies significant gaps between doctors' and patients' understanding of tests.

Expectations of tests

Patients in this study had high expectations of tests. This is in keeping with previous qualitative interviews with patients in general practice waiting rooms in the Netherlands, where patients expected tests to provide diagnostic certainty or proof of good health.(166) Interviews with patients who reported omissions of care also corroborate the high value patients apply to tests.(234) In a systematic review of survey and interview studies that have measured patients' expectations of the benefits and harms of tests and screening, the majority of participants were found to overestimate the benefits and underestimate harms.(167)

Doctors' expectations of tests were more limited, and doctors' decision-making encompassed both medical and non-medical reasons for testing, in keeping with my pre-doctoral qualitative interviews with doctors,(164) and other qualitative research exploring doctors' perspectives of tests.(162, 235-237) This is in accordance with a group of social cognitive theories, known as situativity theories, which conceptualise diagnostic decision making as a social and situated processes, shaped not only by medical knowledge, but by complex, dynamic interactions between doctor, patient and environmental factors.(238) Although previous studies have explored doctors' and patients' expectations of blood tests separately, this study offered the benefits of comparing and contrasting these expectations within individual healthcare encounters, illustrating how mismatched expectations influence the outcomes of testing.

Decision-making

Shared decision-making is a process in which patients and clinicians work together to make decisions based on evidence and patients' preferences, and is widely recognised as 'best practice'.(160, 239) Awareness of the importance of shared decision-making is widespread, yet most research focuses on treatment decisions. This research demonstrated a lack of shared decision-making around

diagnostic testing. Although these findings are based on patient and clinician recall, they are in keeping with research which I co-authored, using video recorded UK general practice consultations, where the majority of blood testing decisions were instigated by doctors, with a lack of information sharing and shared decision-making.(181) Similarly Ford *et al* analysed 212 video-recorded primary care consultations and rated the extent to which patients were involved in decisions as well as measuring patient preferences for decision-making; 91% of decisions about investigations were doctor led, 9% patient led and none of the consultations observed demonstrated shared decision-making around investigations.(182)

Previous ethnographic and focus group research with patients and the general public in a UK hospital setting demonstrated that patients have limited understanding of blood testing, which is seen as a routine, everyday practice which patients routinely accept without questioning.(240) This is in keeping with the perception by patients in this study that blood test decisions were not an area where options or choices were available to them.

Information sharing

This research showed a lack of information sharing, with patients frequently unaware of which tests had been requested and why. This is in keeping with a previous survey by Kljakovic *et al* using patient questionnaires, in which only 19% of patients could name the tests which they had had done.(241) In contrast, Kljakovic et al found that 90% of patients reported that they understood the reasons for blood tests.(241) The higher self-reported levels of understanding in this survey could reflect the different study design, and the different setting, of Australian primary care.

The findings from my research, that most patients were given an assessment of whether their test result was normal or abnormal, but few knew the actual result, is in keeping with conversation analysis research from Pomeranz *et al*, who used

a sample of 33 video-recorded consultations in US ambulatory care, to examine how test results were conveyed to patients; they suggest that offering an assessment only was an example of paternalism, whereas providing test results empowers patients as an independent expert.(242)

Consequences of testing

Tests allow doctors and patients to construct meaning and explanation for patients' symptoms and have value beyond their medical purposes. Kravitz *et al* found that patients perceived tests to have symbolic value, offering a way of validating the patients concerns, demonstrating an interest in the patient and building an empathic connection.(234) This is in line with social constructionist theories, the concept that illness is not only determined by biology and physiology but is socially constructed.(243) Test results have biological meaning, but this meaning is interpreted by patient and doctor in the context of their ideas, concerns and expectations. The concept that tests do more than just passively reveal facts, but that they form part of the social interaction and can generate meaning has been described in the social science literature by Mol, who explored the non-medical consequences of blood sugar measurement on patients.(244) Mol says: '*It fits all too beautifully (and sadly) into a pattern that has often been described in critiques of diagnostic devices: by being put in the position of correcting subjective sensations with objective findings they end up eroding the subjective sensations, or at least, by making them of little relevance in the daily management of (chronic) disease*.' This is in keeping with the finding in this thesis that patients with normal test results may feel that their own subjective experiences were being invalidated. This also resonates with qualitative interviews exploring patients' experiences of diagnostic tests for chronic back pain; patients had high hopes invested in testing and described how negative or normal tests results led to patients feeling delegitimized.(245)

Patients in this study with normal test results who had ongoing symptoms did not report feeling reassured, which contrasted with doctors' accounts. This fits

with the findings of several systematic reviews of quantitative studies measuring patient outcomes following testing, which demonstrate a lack of reassurance from normal tests.(169, 170, 246) Patients wanted more information and explanation of the meaning of normal test results. Improving the communication of normal test results has been shown to lead to greater patient reassurance and reduced reporting of symptoms in a randomised trial in chest pain clinics.(247)

Barriers to shared decision-making

This research identified multiple barriers to shared understanding and shared decision-making, which were categorised into barriers around knowledge, attitudes and systems of testing. Blood tests were perceived by doctors as relatively trivial interventions which were therefore low priority for information sharing. This is in keeping with a recent realist review exploring test ordering in primary care from clinicians' perspectives, which identified a commonly held perception that laboratory tests are relatively inconsequential interventions.(165) In this study it was proposed that doctors, facing high workloads, prioritise efficiency over thoroughness in test-ordering decisions. This 'efficiency-thoroughness trade off' principle described by Hollnagel may underlie the limited information sharing around blood testing.(248) Clinicians interviewed for this thesis described prioritising other aspects of information sharing which were perceived as more important, in the context of time limited consultations. Although workload can be a barrier to shared decision-making, there is evidence that interventions to increase shared decision-making may reduce healthcare utilisation and overall costs;(249) similarly, patient-centred communication has been shown to be associated with lower rates of diagnostic testing.(250)

Another barrier to information sharing identified in this research was the imbalance in knowledge and power between doctors and patients; this has been widely discussed in the shared decision-making literature in relation to treatment decisions, but in a recent systematic review of barriers and facilitators to shared decision-making, none of the 45 articles included explored decisions

pertaining to diagnostic testing.(251, 252) In a qualitative interview study O'Flynn *et al* found that GPs perceived test ordering to be a biomedical decision, that formed part of their professional identify, and was not appropriate or available for shared decision-making.(253) This is in keeping with the perceptions of some GPs in my study that testing was 'still one of those areas' where a more paternalistic approach was justified. O'Flynn *et al* proposed that a shift in perception of medical identity, to enable sharing of power and responsibility, was required to enable more shared decision-making.

Another reason for withholding information about tests which was mentioned by GPs in this study was the wish to protect patients from unnecessary anxiety and uncertainty. Although GPs were concerned that sharing too much information could cause anxiety, patients in this study perceived that improved communication about blood tests could reduce anxiety.

The challenges of communicating uncertainty have been discussed in the wider medical literature.(254) Some have argued that open communication of uncertainty is necessary for true shared decision-making,(255) with some evidence that sharing uncertainty does not reduce trust.(256) Conversely, others have shown that presenting uncertainty may decrease patient understanding in relation to treatment decisions,(257) and explicit expression of uncertainty can be associated with lower perceived technical competence.(258) There is a lack of evidence around communication of uncertainty in relation to tests. One recent trial measured information sharing and public understanding of Covid-19 test results: when participants were given information about tests which incorporated uncertainty, fewer participants interpreted results as definitive, however this information was felt to be more difficult to understand and slightly less 'trustworthy'.(259)

Systems of test result communication were found to be another barrier to shared understanding of blood tests, with uncertainty, variation and a lack protocols for test result communication. This is in keeping with previous surveys of UK general practices which demonstrated that most rely on patients contacting the

practice for their test results, with a lack of fail-safe mechanisms.(158) Qualitative interviews with UK clinical and office staff in primary care demonstrated the complexity, lack of standard protocols and problems with test result communication in primary care,(156) in keeping with this research. Similarly, interviews with US physicians in primary and secondary care demonstrated that many clinicians lacked methods to ensure test results were received and communicated to patients.(260) Studies quantifying failures in test result follow up have been systematically reviewed, with between 6.8% and 62% of laboratory tests reportedly not followed-up in US settings, and no relevant UK research identified.(155) These failures in test result communication can lead to serious lapses in care.(261)

Patients' perspectives have received relatively little attention; focus group discussions with patients about their preferred methods of test communication highlighted patient dissatisfaction with non-clinical staff relaying results,(157, 262) in keeping with my findings. By comparing doctors' and patients' perspectives on a single healthcare encounter, the research presented in this thesis was able to highlight problems which occur when both doctors and patients assume that the other party is responsible for obtaining test results.

Shared decision-making or shared understanding?

Despite the barriers to shared decision-making around testing identified in this study, there are examples in the literature of shared decision-making for tests, mostly in relation to specific tests such as prostate specific antigen (PSA), genetic testing(178, 263) and screening tests.(179, 264) Whilst genetic testing and screening are arguably less relevant to primary care, PSA is commonly tested in primary care. Multiple qualitative studies of PSA testing have explored patients' perspectives.(265) Information resources(266) and decision aids(180) for PSA testing have been developed and systematically reviewed,(177) demonstrating that it is possible to share information on testing in primary care. Although PSA is regarded as something of a 'special case', as testing can lead to patient harms

including invasive further tests and over-diagnosis, arguably similar harms of cascade testing, anxiety and over-diagnosis could occur following inflammatory marker blood tests.

Shared decision-making is recommended by NICE for '*people who use services and their families and carers to choose tests, treatments, management or support packages*', with NICE recommended decision aids for PSA testing, BRCA genetic testing, Down's testing, prenatal screening, HPV testing and colorectal cancer screening.(176) However, there is a lack of discussion in the literature of how this relates to routine blood tests in primary care. Patients in this study did not express a desire to share decisions about every test, but they did express the need to understand the meaning of their test results in the context of their symptoms. Shared decision-making conceptualises the practice of medicine as a series of discrete choices; this does not seem to fit with the experiences and perspectives of patients and clinicians interviewed, who describe tests as only one part of a complex medical and social interaction. A broader model of shared understanding seems to be more relevant to the complexity of primary care diagnosis which rarely involves simple binary choices, but instead involves intuition, ambiguity and uncertainty.(267) Shared understanding of medicine, in contrast to shared decision-making, is a continuous process, described by Richard Lehman as follows: (239)

> *Clinical care does not consist of a series of easily defined take-it-or-leave it choices but is a process of understanding developed and deepened over time. Sharing understanding with patients is a form of dialogue and interaction which cumulatively develops and which effects changes in both parties: it lies at the heart of primary care, and it is essential for kind and effective clinical practice in all specialties.*

6.4. Strengths and limitations of research

6.4.1. Quantitative study

The main strengths of my CPRD research are the size of the study and the setting in primary care, where initial suspicion of disease usually arises. Most previous research focuses on single disease outcomes, and most is based in secondary care; whereas this research examined multiple disease outcomes. This allows exploration of the utility of tests to distinguish between 'healthy' versus 'unhealthy'. This is important as it matches the way the GPs describe using inflammatory markers in practice.(2) The use of a test-consequences graphic based on a nominal population of 1000 tested patients allowed results to be presented in natural frequencies, (184) which are easier for patients and clinicians to interpret,(268) and to aid implementation into practice. Traditional diagnostic accuracy studies only consider the performance of an index test in comparison to a reference standard; however, the use of routine data allowed process outcomes to be measured including rates of consultation, referral and follow-on blood tests, to explore the diagnostic activity following inflammatory marker testing. This is important because the unit cost of these blood tests is low, yet the downstream consequences may be costly.

The large sample size and large numbers of participants with multiple simultaneous blood tests allowed direct comparisons between the accuracy of different inflammatory marker tests. This reduces the potential for selection bias, where tests might perform better for certain disease outcomes due to GPs pre-selecting those at higher risk to have a specific test, for example, preferentially using CRP when an infection is suspected.

Limitations: Reasons for testing

The main limitation is lack of information about the reason for testing; it is not possible from CRPD to determine which tests were done for specific diagnostic

purposes, and which were done as a general rule-out for any relevant underlying disease. To try to gain an indication of the main reasons for testing, the symptoms in the 28 days prior to testing were measured; the frequency of non-specific symptoms in the cohort (tiredness, malaise, dizziness, low mood) suggests that non-specific testing is likely to be common. A limitation of this approach is that symptoms are less likely to be well coded by GPs in CPRD compared to coding of diagnoses. The benefit of this approach is that it reflects real-life clinical practice: although GPs may not have a specific diagnosis in mind when they request inflammatory markers, they need to consider a wide range of possible diagnoses if the test is positive.

Limitations: Reference standard

The reference standard for diagnostic accuracy research is usually defined as the best available method for classifying whether people have the target condition or not. In this study one-year incidence of cancer and autoimmune disease and one-month incidence of infection were used as a proxy measure for presence of disease at the time of testing; thus, the reference standard is dependent on the quality of GP diagnosis and coding of disease. Blood tests are electronically transmitted to the GP records, reducing the risk of missing or erroneous data in the index tests. Rigorous methods were used to develop disease code-lists,(192) but it is possible that there were some omissions, and relevant diseases linked to a raised inflammatory marker apart from infections, autoimmune conditions or cancers could have been missed. Some diseases may have been diagnosed but not coded: though this is likely to be rare for serious conditions such as cancers and autoimmune diseases, it could be more common for minor infections. It is also possible that some diseases remained undiagnosed, or were not diagnosed within one year of the test. This could mean that the numbers of 'false positives' reported in this study could be an over-estimate, as some people with a positive test result could have had relevant disease which was undiagnosed, or uncoded.

The decision to use the time period of one-year for incidence of cancer and autoimmune disease as reference standard was based on evidence of time from symptom onset to cancer diagnosis.(214) To check for a late effect, two-year incidences of cancer and autoimmune disease were also explored. A shorter one-month time period for incident infections compared to cancer and autoimmune conditions was chosen on the basis that patients could have multiple infections within a one-year time period; infections beyond one-month were deemed unlikely to be related to the initial raised inflammatory marker.

There is some risk that the diagnoses made by the clinician were on the basis of the inflammatory marker test result. This is known as incorporation bias, where the index test is part of the reference standard, and generally leads to an overestimation of the sensitivity of a test.(269) For significant diagnoses such as cancer and autoimmune disease this is unlikely to occur, as confirmatory diagnostic testing would usually be expected, but for infections this is more likely to be an issue.

Cardiovascular disease was not included within the reference standard within the definition of 'relevant disease'. Whilst CRP has prognostic value for prediction of future cardiovascular disease,(125) it does not form part of any cardiovascular diagnostic algorithm. Inflammatory marker testing would not therefore be used as a diagnostic test for cardiovascular disease in primary care, and therefore does not fit within the remit of this thesis to explore the diagnostic utility of inflammatory markers.

Linked data including Cancer Registry linkage and ONS death registration data were used to improve outcome ascertainment for the outcomes of cancer and mortality; however, the cohort was not limited to only those eligible for data linkage, lest this introduce bias, as some regions are not linked to the cancer registry. Instead, sensitivity analyses were done to explore the effect of this; minimal differences were seen when analyses were restricted to those eligible for linkage. On reflection, a reasonable alternative approach might have been to limit

the study to patients in England only, to reduce the numbers within the cohort who were ineligible for linkage.

Limitations: Participant selection

Another limitation of the observational cohort design is that patients having blood tests in primary care are a selected group and will be generally 'sicker' than those attending primary care who are not tested, or the general population. This is clearly demonstrated by the fact that the overall disease incidence was lower in the untested comparison group than the tested group. The disease incidence in the test positive group therefore reflects two things: the risk associated with the test result, and risk associated with the clinicians' decision to test. The benefit of having a comparison group was that it allowed the predictive value of the clinicians' decision to test to be quantified. However, this means that caution is needed when interpreting the findings. It is tempting to suggest that these tests could be useful for cancer diagnosis in primary care; however, the results do not necessarily support the use of inflammatory markers to detect cancer in patients who would not otherwise have been tested. This also potentially limits the generalisability of the results, as the findings may not be applicable to different healthcare systems with a higher or lower threshold for inflammatory marker testing. To address this, predictive modelling research is needed, to explore how inflammatory marker test results could be combined with symptoms, signs and other test results to predict future disease outcomes, particularly cancer.

Limitations: diagnostic activity after inflammatory marker testing

The increased rates of consultations in those with 'false positive' inflammatory marker test results compared to 'true negatives' is interesting, but this observed association does not prove causality. One possibility is that some of the 'false

positives' did have relevant disease, which was not diagnosed or coded, or was omitted from my code-list. This could generate an appropriate increase in healthcare usage. Another possibility is that those who are selected for inflammatory marker testing represent high users of healthcare, which is then perpetuated by the abnormal results. On reflection, stronger evidence of causality could therefore have been obtained if I had adjusted for baseline consultation rates.

Limitations: Clinical significance of findings

Initial sample size calculations suggested 80,000 tested and 20,000 untested patients would be sufficient; the CPRD suggested a larger sample size (at no extra cost) could be used to allow extra power for rarer disease outcomes. A consequence of this large sample size is the potential for the study to be overpowered; tiny differences in AUC were detected, which had small p-values, but which are very unlikely to be clinically significant. For example, the overall AUC for relevant disease for CRP and ESR in combination was 0.688, compared to 0.682 for CRP alone. Although the p-value was <0.001, a difference of 0.006 in AUC is very unlikely to be of clinical benefit. This demonstrates the importance of considering results in a clinical context, rather than focussing on measures of statistical significance.

6.4.2. Strengths and Limitations: Qualitative study

The main benefit of this research was the ability to compare doctors' and patients' perspectives on the same healthcare encounter. This allowed direct comparisons which highlighted how mismatches in communication and understanding could play out within a single healthcare encounter and how this could influence patients' experiences.

Strengths and limitations: representativeness

Recruitment was purposive and sampling included a range of patients in relation to age, gender, socioeconomic status and ethnicity; this was important in order to capture a range of views and perspectives. The patient sample had a higher proportion of women (64%); however this is in keeping with the CPRD data which showed that 62% of tested patients were female. All interviews were conducted in Bristol, North Somerset and South Gloucestershire region for practical reasons, although efforts were made to ensure a mix of urban, suburban and rural practices. However, blood testing may have regional variation, and many of those interviewed will have used the same laboratory. This means the findings may not reflect the processes around testing used in other health care systems. However, they do highlight the potential for misunderstanding when information around testing and the communication of results is not shared.

Strengths and limitations: recall bias

The majority of the initial interviews with patients were done at the time of their blood test appointment; this had the benefit of allowing me to capture patients' immediate thoughts and feelings at a time when their tests were foremost in their minds. The second interview allowed me to compare how these perspectives changed after they had received their test results.

The main limitation is that the interviews were based on patients' and GPs' recollection of the healthcare encounter, rather than direct observation of the doctor-patient interaction and communication. This could lead to recall bias; for example, it is possible that GPs may reinterpret their reasons for testing in order to rationalise their decision-making, or to overestimate the information that they shared with patients. However, my experience was that GPs did share openly the limitations in how they communicated blood tests with patients during interviews. The benefit of interviewing patients rather than directly observing consultations, is that it allowed me to find out what patients understand and

retain after a consultation. During the course of my PhD, I was able to supervise an Academic Foundation year 2 doctor, Jess Martin, who analysed the content of video recorded GP consultations from the One-in-a-Million database, to explore communication around blood testing. The findings from these video recorded encounters are in keeping with the findings from this thesis, with limited information and a lack of shared decision-making observed.(181)

Whilst patient interviews were conducted at the time of testing, the time interval between the consultation and GP interviews was longer, due to the practical challenges of trying to arrange interviews with GPs with busy clinical workloads. GPs were asked to ensure they had access to the medical records as an *aide memoire* during GP interviews; nonetheless it is possible that they had incomplete recollection of the specific consultation. Their recollection might therefore partially be based on how they feel they would usually communicate in that clinical scenario.

Strengths and limitations: reflexivity

Overall, my experiences of being both a GP and a researcher was that this helped facilitate my access to GP practices, and may also have facilitated my communication with GPs who seemed comfortable discussing cases with a fellow clinician with shared understanding. The GPs whom I interviewed were aware that I was a fellow clinician, and some were aware of my published research into inflammatory marker testing. In the GP interviews I emphasised that the interviews were non-judgemental, and were focussed on exploring communication around testing, not on scrutinising clinical decision-making. Nonetheless, clinicians who were less confident about inflammatory marker testing may have declined to take part. Other issues arising from my dual role as a clinician and a researcher are discussed further in **section 4.6.**

6.4.3. Strengths and limitations of a mixed methods approach to inflammatory marker testing

The mixed methods approach used in this thesis allowed me to gain very different but complementary perspectives of inflammatory marker testing in primary care. The findings of my pre-doctoral qualitative research(2) informed the questions for my CPRD study; in particular the decision to focus on multiple disease outcomes and to explore the accuracy of inflammatory markers as a 'rule-out' test. Similarly, the findings of my CPRD research informed the questions for my qualitative study; in particular the fact that these tests lead to false positives and false negatives with potential negative consequences meant it was important to explore how information on benefits and limitations of testing was shared with patients (or not) and how patients engage in decision-making.

A limitation to the approach used is that I could not use formal methods to synthesise the results of the qualitative and quantitative research. After undertaking the qualitative interviews, I decided to conduct focus groups with patients and GPs, aiming to bring together the qualitative and quantitative findings of my thesis and explore ways of improving communication and shared understanding of inflammatory markers. I was particularly interested to use these focus groups to explore how patients would respond to information about the uncertainty of tests, and whether they would feel that this information should be shared with them. Unfortunately, due to the exceptional circumstances of the Covid-19 pandemic, it was very challenging to recruit to these focus groups; despite hard work I was only able to recruit three patients from the original qualitative study back for further focus group discussions. Although this could have provided helpful further insights, it was not part of my original doctoral thesis plans. On reflection, a larger piece of work would be needed to generate meaningful insights to improve the shared understanding of tests; this is one of the areas I hope to explore as a post-doctoral researcher.

6.5. Implications for research and clinical practice

6.5.1. Quantitative study

My pre-doctoral qualitative work found that doctors perceive inflammatory markers to be a useful rule-out test for patients with non-specific symptoms.(2) The finding from this thesis that these tests have a low sensitivity contradicts this, demonstrating clearly that they are not suitable as a rule-out test. Instead they are classic Bayesian tests, with a positive test somewhat increasing the chance of disease, though not definitively, and a negative test reducing the chance of disease, but not to zero.

In patients with a raised inflammatory marker, the range of differential diagnoses is wide, which is one explanation for the additional consultations, tests, and referrals. With very high inflammatory marker test results, the risk of disease is higher. Interpretation should take into account the reason for testing and the pre-test likelihood of disease; a negative test in the context of low-risk symptoms reduces disease likelihood further, but with the potential for harm from false-positive tests. False-negative tests may also lead to false reassurance, as patients with normal inflammatory markers are at higher disease risk than untested controls. GPs should not be excessively reassured by a repeat negative inflammatory marker; those with normal repeat tests had higher overall incidence of disease (18.6%) compared to those who did not have repeat tests performed (12.5%). Presumably, this reflects the fact that the GP's decision to repeat the test meant they were more suspicious that the patient had underlying disease.

Testing multiple inflammatory markers does not improve the ability to rule-out disease, but does increase the risk that at least one of the tests will give a false-positive, compared with a strategy of using a single test.

The overall diagnostic utility of all three inflammatory markers is similar, however CRP marginally outperforms ESR and PV for infections. CRP is also cheaper than either ESR or PV (£1.19 for CRP, £3.18 for ESR, £3.18 for PV, source Bristol North Somerset and South Gloucester CCG laboratory costings). It seems sensible therefore to use CRP as the first-line test in most circumstances. Exceptions might include the use of ESR or PV rather than CRP for suspected myeloma, though if there is strong clinical suspicion then direct testing using electrophoresis and Bence Jones protein would be reasonable.

There is no combination of inflammatory markers that can be used as a reliable rule-in or rule-out test strategy. Results and decisions to test must be made in the context of other clinical findings. A negative test in the clinical context of a low-likelihood situation may be sufficient to provide reassurance.

Although inflammatory markers have a moderate predictive value for inflammatory bowel disease (IBD), the AUC for CRP of 0.698 (in a model that includes age and sex) is much lower than for calprotectin, with a published AUC of 0.95,(270) meaning that calprotectin is to be preferred if IBD is under consideration. Similarly, though inflammatory markers have a modest AUC for rheumatoid arthritis, low sensitivities found in the present study are in keeping with previous studies, which have found that 35% to 45% of patients with rheumatoid arthritis have normal inflammatory marker levels at diagnosis.**(56)** National Institute for Health and Care Excellence guidelines therefore recommend referral of patients with clinical evidence of rheumatoid arthritis, even with normal inflammatory marker test results.**(55)** It is therefore hard to see any benefits from inflammatory marker testing where rheumatoid arthritis is suspected diagnostically, though it may have a useful role in disease monitoring.

Cancer is worthy of special consideration, as current UK NICE guidelines recommend urgent cancer referral for any patient with a risk of cancer of 3% or higher,(65) with studies of patient preferences suggesting an even lower threshold of 1%.(271) Inflammatory markers are not recognised within current guidelines for cancer diagnosis,(65) with the exception of myeloma, where first

line tests include ESR or PV. With overall PPVs of 3.53%, inflammatory markers may therefore have some role in early diagnosis of cancer.

Interpretation of inflammatory marker test results must take into account the reasons for testing: if there is a clinically obvious explanation for raised inflammatory markers from history and examination, then further investigations for cancer would not usually be appropriate. Women under 60 and men under 50 with raised inflammatory markers have a risk of cancer below the 3% threshold, and in the absence of other risk factors, further investigations for cancers would not usually be warranted. For older patients with unexplained raised inflammatory markers, these findings support a strategy of repeat testing, with lower cancer incidence in those for whom the test returns to normal, and higher cancer incidence in those with rising inflammatory markers. Further investigations for cancer should be considered in patients with persistent unexplained raised inflammatory markers, particularly in older men, who are at highest risk. With significantly raised inflammatory markers, especially if accompanied by 'low-risk but not no-risk' cancer symptoms or signs, urgent investigation or referral may be appropriate, without repeat testing. The range of possible cancers is wide, so the choice of further investigations will vary depending on the clinical history and examination findings; recently introduced multidisciplinary diagnostic centres for patients with non-specific but concerning symptoms may be appropriate if a clear source cannot be found. Excluding myeloma is not sufficient, as this only contributes a small proportion of the cancers diagnosed in the raised inflammatory marker cohort.

Although the PPVs from a raised inflammatory marker are clinically significant, inflammatory markers have poor sensitivity, so cannot be used to rule-out cancer. The clinical usefulness of these finding are therefore in guiding the clinical management of a patient with a raised inflammatory marker. Inflammatory markers should be used judiciously for possible cancer, taking into account the risks of false positives, which may generate anxiety, and false negatives, which may generate inappropriate reassurance. Further research is

needed to explore how inflammatory marker blood test results, combined with other common blood tests, such as platelet, haemoglobin and calcium levels, alongside symptoms and signs, could generate prediction models with increased accuracy, to improve the early diagnosis of cancer in primary care.

Although a clear association between inflammatory markers and one-year mortality was identified, the clinical utility of this finding is less clear-cut. It seems unlikely that clinicians would test inflammatory markers for the purpose of mortality prediction as this would be likely to cause worry and anxiety to patients, particularly as there is a lack of evidence-based interventions available to target those identified at high risk.(272) However, general practitioners are already required to identify patients who are frail,(273) and inflammatory marker tests are commonly performed for many other reasons. Inflammatory marker test results, when available, may therefore add useful information to improve prediction of mortality and assessment of frailty in primary care. General practitioners should interpret raised inflammatory markers within the wider clinical context; where the cause of inflammation is identifiable and treatable, mortality risks should not cause undue alarm. However, clinicians should consider whether older patients with persistently raised inflammatory markers are reaching the end of life.

6.5.2. **Qualitative study**

The qualitative research findings have important implications for clinical practice, and are relevant, not only for inflammatory marker testing but blood testing more widely. The lack of shared understanding is important, not only for the ideals of shared decision-making, but also for the more fundamental principles of informed consent. The General Medical Council (GMC) duties of a doctor state '*you must give patients the information they want or need about the purpose of any proposed investigation*'.(173) The fact that less than half of the

patients interviewed (11 out of 28) perceived that they knew the reason for testing is at odds with this.

Shared understanding is not only important for informed consent, but also for doctors and patients to create meaning from the test results. Shared understanding and explanation are important, not only for patients with abnormal test results; patients in this study with normal test results also needed explanation and understanding of their test results. There are several potential harms from a lack of shared understanding. Firstly, if patients are offered a false promise of certainty from tests, they can feel frustrated, or even delegitimised by normal test results. Secondly by imbuing test results with too much weight, patients may be falsely reassured by normal results (exemplified by the statement from one patient that '*obviously there's no cancers going on*'). If normal results are not accompanied with appropriate safety netting this could potentially delay appropriate health seeking. Future research into diagnostic tests should measure not only the diagnostic value of tests but also consider patient reported outcomes including cognitive, emotional, social and behavioural outcomes of testing.(274)

Although doctors were concerned that sharing too much information could generate anxiety, patient interviews did not show any evidence for this; there was, however, evidence that a lack of information sharing could cause anxiety. Patients did not want to be overloaded with technical information or medical terminology about their tests, but they did want information in lay terms about why tests were being done, and what the results would mean for their health. Clinicians who proactively shared their expectations of tests before the results were available, were able to improve shared understanding of tests.

Multiple barriers to shared understanding were identified. One of these was a perception from doctors that sharing information about tests was low priority and that patients did not want this information. The findings of this research challenge this assumption.

Systems of testing were a barrier to effective communication, and a lack of clear protocols for sharing test results made it challenging for patients to navigate these systems. These findings highlight the risks of clinicians assuming patients will proactively seek out their test results by making contact with the GP surgery. This is important because failures in test result communication, which are identified in this thesis, are a potential area for patient harm and litigation. Further research and quality improvement for systems of blood test communication is needed. This should engage key stakeholders including patients, doctors, and the wider healthcare team. Co-production methods could be used to collaboratively develop and improve test communication. This should include fail-safe mechanisms to ensure results are returned to patients to prevent harms.(275) It should also consider the importance of improving the experience of patients whose test results are normal. Whilst new opportunities such as text message systems, and online access to test results have potential to enhance communication, patients' perspectives should be sought and potential for causing anxiety should be carefully considered.

A lack of resources for information sharing was another barrier to communication. Future research is needed to generate resources and interventions to improve shared understanding of tests; these should be designed with input from both clinicians and patients.

Time pressures within 10-minute GP consultation and GP workload were perceived to be barriers to improved communication. Although this is a potential challenge, previous studies have found that improved patient centred communication is associated with lower rates of diagnostic testing.(250) Improving the communication and shared understanding could therefore be important not only for patient satisfaction but could also potentially reduce workload by reducing testing rates and reducing re-attendance following blood tests.

6.6. Synthesis and reflections

The quantitative results from my thesis illustrate and quantify the complexity and uncertainty inherent in inflammatory marker testing. Far from providing the clear-cut answers which patients expect, these tests have modest diagnostic accuracy which can neither be used to rule-in nor rule-out, yet may offer 'clues' to possible serious diagnoses. This is very different to patients' expectations and understanding of these tests as revealed by the qualitative component of this thesis. GPs are aware of the limitations of tests yet seem to be reluctant to share this with patients, with multiple barriers to shared understanding identified. Patients needed to understand the meaning of their results, in the context of their symptoms; without this 'normal' results did not always provide the reassurance that GPs were expecting. This is particularly important given that inflammatory markers are frequently used as a tool for ruling-out disease and reassuring patients.(2) The findings of this thesis contradict this - inflammatory markers do not have the necessary properties for a 'rule-out' test, nor do these tests, in themselves, provide reassurance to patients, without explanation and understanding. Inflammatory marker testing therefore has limited diagnostic utility and limited non-medical benefits for patients. Yet highlighting the limitations of tests does not mean that the tests have no place in clinical practice. My interviews with patients and GPs highlighted cases where doctors made carefully balanced clinical decisions to test and shared this with patients, consequently the test results helped clarify the diagnosis and provide patient satisfaction and reassurance. This demonstrates the importance of triangulation of qualitative and quantitative methodologies, to uncover the stories within the data, and the underlying meanings created from tests.

The contradictions between patients' expectation of tests and the reality of test result interpretation raise many unanswered questions. Given this complexity for a single test, to what extent is shared decision-making, for a battery of blood tests, possible? Can these uncertainties be meaningfully shared with patients within a 10-minute consultation, and 'should' they? Is shared decision-making

the 'right' model for decisions on testing, or is a different approach warranted? There is a tendency to assume that all shared decision-making is 'good', and all paternalism is 'bad', yet in my view the reality is more nuanced. Doctors have expertise and make decisions using clinical experience and intuition in the face of diagnostic ambiguity and uncertainty. Patients speak about 'trusting' their doctor to make the right decisions, and do not perceive blood testing to be an area where options are open to them. Shared decision-making requires sharing uncertainty, whereas patients generally perceive tests as a solution to resolving uncertainty.

I had an opportunity to begin to explore some of these issues through meetings with my PPI group where I shared the findings of my quantitative and qualitative research. I summarised my quantitative results in a single infographic in the form of a modified cates plot, to demonstrate the outcomes of inflammatory marker testing outcomes in natural frequencies (see **Figure 10**). Participants in the meeting were surprised by the findings, which did not match their expectations and understandings of tests. Whilst they felt strongly that communication about testing should be improved, they were not strongly supportive of the principles of shared decision-making for blood tests and were unsure whether sharing information on the limitations of tests would be helpful for patients. In general, the PPI panel felt that they would value more information after their blood tests were done, rather than before. This highlights the risks of assuming that sharing decision-making is the 'right' thing to do, without exploring the consequences and patients' perspectives.

A range of different ideas for improving understanding of tests were explored: one suggestion from the PPI group was that posters in GP practices, similar to the 'ask three questions' campaign,(276) could help empower patients to find out about their blood tests. Patient information leaflets on blood tests were discussed; the PPI group highlighted that because these provide general information, they were unlikely to answer patients' questions about the meaning of test results for their health, in the context of their symptoms. The PPI group

were much more interested in the idea of providing personalised written information at the end of a consultation, summarising briefly which tests had been requested and why, and how to obtain test results. I am currently engaging with colleagues in Bristol to explore how this could be incorporated into the Consultation Open And Closed (COAC) Study.(277)

My own experiences as a clinician have been influenced strongly by my research into inflammatory marker testing, and I now only use these tests in a limited number of circumstances – most commonly for suspected polymyalgia. My communication with patients has also been strongly influenced by my qualitative research, yet despite my desire to improve communication of tests, I often find it challenging to achieve a shared understanding. I also often find myself reluctant to disabuse patients of the expectation that tests will provide clear-cut 'answers'. I find myself using easy phrases when arranging tests like "we'll make sure we're not missing anything", even though I know that this is over-optimistic. Changing the nature of these conversations is challenging, and further research and education is needed to try to improve communication and promote a shared understanding. I have made significant changes to the ways in which I share test results, and I make an effort to contact patients to discuss the meaning of their test results in the context of their symptoms.

Overall, management of uncertainty seems to be the overarching theme highlighted by the interplay between the qualitative and quantitative aspects of this thesis; uncertainty as a reason for testing, uncertainty in interpreting test results, patient uncertainty as to why tests were done and how to get results back. Some of this uncertainty can be reduced by research and quality improvement, for example improving systems of test result communication could reduce patient uncertainty around test results. However, if patients are accustomed to the perception that blood tests will provide clear cut answers, then this research could also increase uncertainty for patients, by demonstrating that inflammatory markers, like all tests, lead to false positives and false negatives. Communicating this uncertainty is a fundamental challenge which has

relevance beyond the realm of diagnostic testing, indeed tolerating uncertainty was suggested by Simpkin and Schwartzstein as being 'the next medical revolution'.(255) In the article they state '*although physicians are rationally aware when uncertainty exists, the culture of medicine evinces a deep-rooted unwillingness to acknowledge and embrace it*'.(255) This culture of medicine may in part underlie doctors unwillingness to share the uncertainty inherent in inflammatory marker testing.

6.7. **Impact**

Since publication, the research presented in this thesis has generated significant interest and impact; with two papers(278, 279) published in the British Journal of General Practice (BJGP) reaching 5th and 3rd place respectively in the BJGP top-10 research papers of 2019. Two GP education providers, NB Medical Education and Red Whale, included these papers in their popular GP Update courses and in the Red Whale's 'best of 2019' webinar. Locally this was written up by the Applied Research Collaborative (ARC) West as a 'BITE' summary.(280) This generated significant interest from Bristol North Somerset and South Gloucestershire (BNSSG) Clinical Commissioning Group (CCG). I was invited to join the CCGs Clinical Pathology Group and as part of this group I developed an intervention to optimise inflammatory marker testing locally, using guidelines, educational outreach and popups built into the electronic test ordering systems. This allowed me to gain experience of implementation before my PhD had finished and build collaborations and networks with clinicians and commissioners. The intervention was iteratively developed following presentations and feedback at BNSSG GP membership forums, and nurse practitioner forums and was endorsed at CCG Locality Leadership Groups. A guideline for inflammatory marker testing was developed in collaboration with GPs from the Clinical Pathology Group and secondary care physicians from gastroenterology, haematology, care of the elderly and rheumatology across the

three local hospital trusts (see **Appendix E**). Pop-ups were implemented in the GP electronic test ordering systems with links to these guidelines on the CCG website (see **Appendix F**). Plasma viscosity and ESR were removed from the 'front screen' of the electronic blood test ordering system, and a popup requiring clinicians to choose a valid indication for testing for PV and ESR was introduced, based on feedback from secondary care clinicians. If a valid indication was not chosen, the blood test automatically defaulted to CRP as first line test.

These popups were implemented in January 2020, with a plan to use an interrupted time series analysis to measure the impact on rates of inflammatory marker testing. Unfortunately, the Covid-19 pandemic in March 2020 led to a near complete cessation of all primary care blood testing, which meant that the planned interrupted time series analysis was no longer possible, and a simple descriptive analysis showing rates of testing over time was presented to the CCG instead. However, the pandemic made my research on diagnostic testing very topical, leading to opportunities for dissemination and impact (see **Box 1**).

Box 1: Wider impact - reflections on diagnostic research in the era of Covid-19

During the final year of my Doctoral Research Fellowship the Covid-19 pandemic hit the UK, and my interest in diagnostic test accuracy suddenly became highly topical. Very early in the Covid-19 pandemic I was concerned that the government and media portrayed Covid-19 PCR tests as offering clear-cut binary answers, and that there was lack of clinical interpretation of test results. I discussed this with colleagues in the Royal College of General Practitioners (RCGP) Overdiagnosis Group, who shared my view that a Bayesian interpretation of tests was important. At this time, during the peak of the first wave, testing was very limited and was reserved for people in hospital with a high pre-test probability of disease, increasing the likelihood of false negative test results. I wrote a paper on interpreting a Covid-19 test in the

BMJ.(281) This paper has an Altmetric score of 4244 (17th highest Altmetric score for a BMJ article ever). I also wrote a companion piece for The Conversation,(282) recorded two podcasts on Covid-19 testing for the BMJ(283) and the RCGP,(284) and was interviewed for BBC Radio 4's More or Less.(285) This led to a significant amount of media interest including coverage in BBC News, New Scientist, World Economic Forum, the Guardian and the Mail Online. The BMJ article has also been adapted for Science in School, the European journal for science teachers. I was subsequently asked by the BMJ to write a second follow up article on interpreting SARS-CoV-2 antibody tests.(286) More recently I have been invited to co-author an article on interpreting a lateral flow SARS-CoV-2 antigen test, which has been submitted to the BMJ and is currently under review.(287)

The Covid-19 pandemic illustrated just how important and challenging it is to communicate the uncertainty inherent in all tests with patients and public. This has been particularly challenging within a social media culture which tends to polarise debates. I hope that the infographics and interactive tools which I helped develop for the BMJ articles are helpful for clinicians and researchers to communicate evidence on test accuracy to a lay audience, not only for COVID-19 testing but also diagnostic testing more generally. I have adapted these infographics for this thesis to summarise evidence on inflammatory marker test accuracy (see **Figure 10**). This experience has helped me to move beyond my limited sphere as a researcher with expertise in inflammatory marker testing and develop my expertise in broader primary care diagnostic test accuracy research.

6.8. Conclusions

The overall aim of this thesis was to explore the diagnostic utility and clinical practice of inflammatory marker testing in primary care, including how results are shared with patients.

My clinical experience and previous qualitative research helped me to frame research objectives of relevance to routine clinical practice, by examining the outcome of 'any relevant disease'.

The findings in this thesis will be of interest to primary care clinicians wanting to know when (and when not) to use inflammatory markers, how to interpret the results and how to communicate this with patients. I have shown that, contrary to GPs' perceptions, these tests are not a useful 'rule-out' test, in fact they miss around half of relevant disease in primary care. Testing more than one inflammatory marker simultaneously does not increase diagnostic accuracy. Whilst not a useful rule-out test, a raised inflammatory marker should trigger consideration of a possible underlying malignancy. Although the tests are often used for reassurance, patients with ongoing symptoms perceived that normal results were unhelpful.

As well as having important implications for primary care clinicians, the findings in this thesis will also be of interest to commissioners wanting to reduce unwarranted variation in testing. Whilst the costs of individual inflammatory marker tests are low, the additional follow up blood tests, GP appointments and referrals are costly. Simple measures such as avoiding using inflammatory markers as a non-specific rule-out, and avoiding the use of two simultaneous inflammatory markers, could reduce unnecessary testing. Authors of guidelines for investigating irritable bowel syndrome, tiredness and dementia should reconsider their recommendations to use inflammatory marker testing to 'rule-out' any relevant disease.

Finally, and most importantly, the findings in this thesis are significant for patients. The quantitative results of this thesis are important to help ensure patients get inflammatory marker tests done when they need them (and avoid them when they don't). The qualitative results in this thesis identify barriers to communication and shared understanding of tests which is important for patients having blood testing in primary care. Further work is needed to move beyond identifying problems with test communication, towards developing

solutions. The issues raised in the qualitative component of this thesis are relevant not only to inflammatory marker tests, but to the wider context of blood test communication and shared understanding between health professionals and patients.

At a more philosophical level, this work raises unanswered questions about shared decision-making and sharing uncertainty. Shared decision-making might be accepted as best practice, but it did not reflect the reality of doctors' and patients' accounts of testing. Patients expressed a desire, not for shared decision-making, but for a shared understanding of tests. Uncertainty is a core part of medicine, and is especially relevant in primary care, yet this uncertainty is rarely shared openly with patients.

My experience over the past five years of generating and exploring my own research questions has cemented my enthusiasm for a career in academic primary care. The methods used in this thesis will be relevant for exploring the use of other blood tests in primary care, and I hope will give me the grounding for a fruitful future career in primary care diagnostics.

Appendix A: Autoimmune diseases and infections included in analysis

Autoimmune diseases*

Inflammatory arthritis	Rheumatoid arthritis
	Seronegative arthritis, including ankylosing spondylitis, psoriatic arthritis and reactive arthritis
Diseases of the endocrine system	Type 1 diabetes mellitus
	Primary adrenocortical insufficiency (Addison's disease) and other autoimmune polyglandular syndromes
Polymyalgia rheumatica	Including temporal arteritis, polymyalgia rheumatica and giant cell arteritis
Vasculitis	Including polyarteritis nodosa, Kawasaki, Churg-Strauss, thrombotic microangiopathy, Wegener's granulomatosis, Henoch-Schoenlein purpura, allergic granulomatosis, Bechet's, cryoglobulinaemic vasculitis, microscopic polyangiitis, granulomatosis with polyangiitis, eosinophilic granulomatosis with polyangiitis, necrotising arteritis, necrotising polyangiitis, chronic granulomatous arteritis, Takayasu's arteritis
Connective tissue disorders	Systemic lupus erythematosus
	Systemic sclerosis (scleroderma)
	Polymyositis/dermatomyositis
Sarcoidosis	
Diseases of digestive system	Inflammatory bowel disease
	Hepatobiliary diseases including primary biliary cirrhosis, primary sclerosing cholangitis, autoimmune hepatitis, autoimmune cholangitis, autoimmune pancreatitis.
Diseases of the skin	Including bullous skin diseases; pemphigus and pemphigoid
Haematological diseases	Including pernicious anaemia, autoimmune haemolytic anaemia, idiopathic thrombocytopenic purpura, antiphospholipid syndrome, immune neutropenia, acquired aplastic anaemia, paroxysmal nocturnal haemoglobinuria PNH, acquired haemophilia,

	thrombotic thrombocytopenic purpura, cold agglutin disease
Diseases of the cardiovascular system	Including rheumatic heart disease, autoimmune cardiomyopathy, autoimmune myocarditis
Diseases of the respiratory system	Including idiopathic pulmonary fibrosis, cryptogenic organising pneumonia, bronchiolitis obliterans.
Diseases of the renal system	Including nephrotic syndrome, IgA nephropathy/Berger's disease, glomerulonephritis, SLE nephritis anti-glomerular basement membrane disease

*Autoimmune thyroid disease, coeliac disease and psoriasis were excluded as it was felt to be clinically implausible that inflammatory markers would be used for diagnosis of these conditions, and because preliminary analysis showed no association of these conditions with a raised inflammatory marker.

Infections

Lower respiratory tract infections	Including pneumonia, pulmonary TB, infective exacerbation COPD/asthma, empyema, lung abscess
Bacterial upper respiratory	Including mastoiditis, streptococcal pharyngitis, otitis media, tracheitis, tonsillitis, sinusitis
Viral upper respiratory	Including laryngitis, pharyngitis, influenza, common cold
Lower urinary tract	Urinary tract infections (excluding asymptomatic bacteriuria)
Upper urinary tract	Pyelonephritis
Gastrointestinal	Gastroenteritis, colitis, enteritis of likely bacterial origin
Diverticulitis	Bacterial infections of diverticular disease, diverticular abscess
Peritoneal infection	Including peritoneal abscess, peritonitis
Hepatobiliary infections	Hepatitis, liver abscess, cholecystitis, cholangitis, empyema of the gallbladder
Bone infections	Osteomyelitis, septic arthritis, pyogenic arthritis
Muscle	Including infective myositis, necrotising fasciitis, gas gangrene, muscle abscess
Skin infections	Cellulitis, erysipelas (excluding localised superficial infections)
Genital infections	Including pelvic inflammatory disease, salpingitis, endometritis, oophoritis, cervicitis, epididymitis, orchitis, prostatitis (excluding superficial infections e.g. vulvovaginitis, balanitis)
Puerperal infections	Including mastitis, endometritis, puerperal sepsis
Cardiovascular infections	Including endocarditis, myocarditis, pericarditis
Neurological infections	Including meningitis, encephalitis, cranial abscess
Other	Any other infections not specified above

Appendix B: ESR upper limit of normal

	Men	Women
<40	11	14
40-49	12	15
50-59	14	17
60-69	14	18
70-79	20	22
>80	20	23

Derived using the CPRD mean upper limit of normal stratified by age and gender, rounded to the nearest integer, in mm/hr

Appendix C: Topic guide

Initial interview with patient participants	
Suggested flow / questions	**Suggested prompts, if needed**
1. Background. Can you tell me a bit about yourself?	• Work life: What do you do for a living? • Home life: Who is at home with you?
2. Blood test. What prompted you to book a blood test today?	• Did a GP or nurse ask you to book a blood test? • Did you get a call or letter from the practice to book bloods? • Do you know why?
2a. What prompted you to see your GP? (If blood tests requested following a GP appointment)	• What were you expecting or hoping for from the GP appointment? • Were you hoping or expecting to have blood tests?
2b. Do you know which chronic conditions are being monitored? (If blood tests being done for routine monitoring)	• Do you know why?
3: Blood tests. What do you know about the blood tests that your GP/nurse requested?	• Do you know which blood tests you are having/have had? • Do you know what these blood tests can pick up? • If inflammatory markers mentioned, then probe further *'what does 'inflammatory marker' mean?'*
4: Communication. What did your GP tell you about your blood tests?	• Did they tell you why they wanted to do blood tests? • Did they give you any options or choices about testing? • Was the decision to test mostly your decision, mostly your GPs decision, or a shared decision? • Did they discuss any benefits or limitations of blood tests? • Did they tell you what the test results might show? • Did they explain how to get the results?

	• Was there anything else about your blood tests that you wanted to know? • Any sources of information you use to get information about tests?
5: Expectations. What do you expect will happen next?	• Do you know how to get your test results? • Do you know when to get your test results? • How would you like to get your results? • What do you think your test results will tell you? • Do you think the test results will change anything for you?
6: Phlebotomy appointment (if relevant). What was your experience of having your blood test done?	• How easy or difficult was it for you to book an appointment? • How easy or difficult was it for you to have your blood test taken? • Did you discuss your blood tests with the phlebotomist/nurse who took the blood?
7: Suggestions for improvement. Is there anything about your experience of blood testing which could be improved?	• Any suggestions to help doctors and nurses communicate? • Any suggestions to help improve the systems? • Any suggestions for patients?
8: Is there anything else about your blood tests you would like to say that we haven't mentioned yet?	
Second interview with patients after blood test results available	
1. Have you received your test results?	If no test results received: • How do you feel about that? • Do you know what happens next? • Would you be happy to reschedule a follow up interview?
2. What was your experience of getting your test results?	• How did you get your test results? (telephone, internet, face-to-face?) • Did the practice contact you, or did you call them? • Who gave you the test results? (GP, nurse, receptionist?)

	• Do you know which blood tests you have had? • Were you given the test results in numbers or just told that they were raised or normal? • What explanation of the meaning of the test results did you receive? • *If test results were abnormal* – what further instructions were you given? (e.g., book appointment with GP, book repeat blood test) • Did you use any other sources of information to find out about your tests? (ego websites, discussions with friends or family)
3. What did your blood test results mean for you?	• Did the results change anything for you? • Do you know what the results mean for your health? • Do you know what happens next?
4. Do you have any suggestions for how communication about testing could be improved?	• What do you want to know about tests? • How do you want to receive information about tests? • Are there any resources you use to find out about tests? • Are there any resources you would like to have about tests?
5. Is there anything else about your blood tests you would like to say that we haven't mentioned yet?	
GP interview	
1. Background. Can you tell me a bit about yourself?	• Years' experience • Type of practice • Role in the practice (partner/salaried/locum)
2. Patient background. Can you tell me what you remember about your consultation with patient X?	• Why did they come to the doctors? • What do you think they were expecting? • What do you think they were worried about?
3. Choice of test. What prompted you to	• Why did you choose to check inflammatory markers?

check bloods on this patient?	• What were you looking for? • How do you decide which bloods to check?
4. What were your expectations of the tests?	• What did you think the results would show?
5. Test results. What did the test results show you?	• Did they change your management? • What will you do next?
6. Patient perspectives. What do you think the patient understands about testing?	• Do you think that patient wanted tests done? • Was the decision to do tests mostly your decision, mostly the patients' decision, or a shared decision? • Do you think the patient knows which tests were done? • What do you think the patient understands about the tests? • What do you think the patient expected from the tests?
7. Communication around testing. What did you explain to the patient?	• What did you explain to the patient about the blood tests you were doing and why? • How will/did you explain the test result? • How do you decide how much information to share with patients? • Are there any resources you use to explain tests to patients? • Is there anything which you would find useful to improve communication around blood testing?
8. Systems of testing. How do systems of testing help or hinder communication with patients?	• Any issues around how the test results are communicated to patients? (text/email/phone calls) • Any issues with communication between hospital/laboratory and primary care of test results? • Any suggestions for improvement to the systems of testing?
9. Is there anything you would like to say that we haven't mentioned yet?	

Appendix D: Coding framework

Number	Name	Description
1 Background		Background information about the patient or GP
1.1	Personal	Background information about the person being interviewed
1.2	GP practice	Information about the GP practice or practice area
2 Why?		Reasons for testing and expectations of tests
2.1	Symptoms	Where doctors or patients describe the symptoms which triggered testing, but not the diagnostic reasoning
2.2	Co-morbidities	Where doctors or patients describe the other pre-existing conditions, which might be linked to testing
2.3	Diagnosis – specific	Tests being done to check for a specific condition
2.4	Diagnosis – general	To check for a number of conditions or more general range of conditions (e.g., 'infections' or 'autoimmune')
2.5	Screening or 'fishing'	Tests done as a general 'screen' or 'fishing' without any particular diagnostic rationale
2.6	Avoiding 'missing something'	When doctors or patients express that they are testing to avoid 'missing something' (might be specific or general) or as a 'rule-out' or to make sure 'everything's ok'
2.7	Answers	Expectations that tests will provide answers
2.8	Magic	Perceptions that tests are powerful, mysterious, or can 'reveal' something special
2.9	Monitoring	To monitor a known diagnosis or drug monitoring
2.10	Secondary care driven	Requests or perceived requests from secondary care
2.11	Patient pressure	Patient requests or perceived pressure from patients including where patient has had pressure from other sources (family/media for

		example)
2.12	Unknown/unsure	If patient or doctor does not know or is not sure why the bloods have been requested or what to expect
3 Which tests		
3.1	'Everything'	When patient express an expectation that bloods will provide a broad 'screen'/'MOT' or check most things.
3.2	Inflammatory markers	When inflammatory markers, CRP, plasma viscosity or ESR are mentioned specifically
3.3	Other specific tests	Include all specific named tests such as blood count, cholesterol, liver, kidney, vit D etc.
3.4	'Batteries'	When doctors or patients mention groups of tests done together (differs from 'everything' as includes a combination of specific tests tailored to symptoms e.g., 'tiredness screen')
3.5	Unknown/unsure	When participants are not sure which tests are being done
4 Phlebotomy		
4.1	Phlebotomy experience	Any description of the experience of blood taking
4.2	Phlebotomy communication	Any description of the communication at the time of blood taking
5 Test results		
5.1	All clear / normal	Results described as 'normal' 'satisfactory' 'fine' 'OK'
5.2	Stable	Results described as 'stable' or 'unchanged'
5.3	Borderline	Results described as 'slightly' raised or 'borderline'
5.4	Abnormal	Results described as 'high', 'low', 'abnormal'
5.5	Numbers	Results described as actual numbers
5.6	Written results	When test results are printed or given in writing

5.7	Unknown/uncertain	Results not known or unclear/uncertain
5.8	Unexpected results	Incidental or unexpected abnormalities in test results
5.9	Meaning	Explanations of what results mean in terms of a persons' health
6 Next steps:		
6.1	Repeat testing	Repeating the same test(s) again
6.2	Further testing	Doing new/extra tests as follow on (either bloods or x-rays/scans etc.)
6.3	Follow up GP appointment	Seeing the GP for a face to face or telephone follow up
6.4	Treatment	Starting a new tablet/treatment/change in management
6.5	Referral	Referring onwards to hospital/another doctor/clinician
6.6	Watchful waiting	'keeping an eye' on things, with an active plan of follow up
6.7	No change in management	When there are no planned next steps, or patient is continuing as before
6.8	Unknown/uncertain	Where doctor or patient is unclear, unsure or does not know what is supposed to happen next
7 Knowledge/understanding		
7.1	Knowledge around testing	Code under this heading what patients and doctors know about tests generally
7.2	Lack of knowledge	When doctors or patients discuss their lack of knowledge about tests
7.3	Guesswork/tacit knowledge	'I guess' 'I imagine'. Knowledge that people have 'picked up' through experience but not been taught
7.4	Misinformation	Any examples where patients make incorrect statements of knowledge or fact
7.5	Inflammation	Capture patient descriptions of what knowledge and understanding they have of the term 'inflammation' or 'inflammatory

		markers' here
7.6	Limitations of tests	Where GPs or patients describe knowledge or understanding of the limitations of blood tests
8 Communication		
8.1	Information sharing	Discussions or descriptions of how doctors share information with patients
8.2	Decision-making	Discussions about how decisions are made – either shared decision-making or paternalistic or patient centered decisions
8.3	Openness	Perceived 'openness' of communication – how willing doctors are to communicate freely and openly with patients or perceived lack of open communication – including withholding information in order to 'protect' patients
8.4	Medical jargon	Use of medical jargon or medical terminology rather than using lay language
8.5	Asking questions	Where patients ask questions, or conversely where they describe finding it difficult to ask questions when communicating with doctors
8.6	Information overload	Where doctors or patients discuss challenges of getting too much information or getting confused or overloaded by information
8.7	Not enough information	Where doctors or patients discuss challenges of not receiving enough information from medical professionals about tests
8.8	Retaining information	Where doctors or patients discuss challenges of remembering what was said during a consultation or forget what they have been told
Information		
9.1	Sources of information	Where doctors or patients get information about tests, or where they would like to get information from
9.2	Websites	When people discuss using online sources of information
9.3	Leaflets	When people discuss using leaflets or written sources of information

9.4	Family/friends	When people ask family or friends for information about testing
9.5	Experience	When people gain information about tests from previous experience
10 Psychology of testing		
10.1	Anxiety	Patient or doctor feelings of anxiety/worry or concern about testing or test results
10.2	Reassurance	Patient or doctor feelings of reassurance
10.3	Uncertainty	People expressing a *feeling* of uncertainty
10.4	Trust	Feeling of trust in doctor patient relationships
10.5	Frustration	Feeling of frustration, annoyance or even anger
10.6	Guilt, internalised responsibility, blame	Feelings of guilt or blame or feeling responsible or that you 'should have' done something
10.7	Disappointment	Feelings of disappointment, for example that tests have not fulfilled expectations
11 Systems of testing – How?		Any discussion of the systems around test ordering and methods test communication
11.1	Text messages	Any discussion of the use of texts as a system for communicating results
11.2	Online results	Any discussion of patient access to results online or via email
11.3	Phone calls	Any discussion of using telephone calls as a system for communicating
11.4	Face to face	When a patient gets test results face to face from a clinician
11.4	Unclear systems	Discussions where there seems to be a lack of clarity or uncertainty about how to get test results
11.5	Assumptions	Systems which assume that things will happen, or assume that someone else is taking responsibility for test communication
11.6	Patient records	How discussions are recorded in the patient

		records/GPs relationship with the patient records
11.7	Electronic test ordering systems	Any discussions of the electronic ('ICE') test ordering systems
12 Who?		Discussion about who communicates test results or information about testing
12.1	Doctors	When doctors communicate results to patients
12.2	Receptionists	When results are communicated via receptionists or admin staff
12.3	Nurses/AHPs	When test results or information about tests comes from nurses, health care assistants or any other members of the health care team
12.4	Practice initiated communication	Where the practice or clinical team proactively contact the patient to inform them of the test results
12.5	Patient initiated communication	Where the patient contacts the GP practice to obtain their results
13 Wider system issues		Discussions about systems issues which are not specific to testing
13.1	Time pressure	Discussions about consultation length or time pressures
13.2	Access	When challenges with accessing primary care are mentioned (e.g., difficulty getting an appointment or getting through on telephone)
13.3	Primary-secondary care interface	Systems of interface between primary and secondary care
13.4	Continuity	Discussions about continuity of care – seeing the same GP or same clinician
13.5	Other?	Need to see if other systems issues are arising in other interviews?
14 Patient engagement		
14.1	Active/engaged	Patients who are active, engaged, want to understand and be involved in decisions about

		their care
14.2	Passive	Patients who are passive, prefer to trust their doctor, don't feel that they can be involved in decisions about their care
Other		**Assorted – these were added in an iterative way during the course of the analysis as frequently occurring themes which were not adequately covered elsewhere**
15.1	Unanswered questions	Where patients have unanswered questions about their tests or about their medical issues which have not been addressed
15.2	Ongoing symptoms	Where patients describe ongoing symptoms, which persist after testing
15.3	Waiting	Discussions of the experience of waiting for test results to come back
15.4	Study impact	Any examples where patient or GPs seem to have been influenced by their participation in the study
15.5	Imagery	Use this code to capture any interesting imagery/metaphors/powerful quotes that might want to explore further?

Appendix E: BNSSG Inflammatory Marker Testing Guideline

Available at: *https://remedy.bnssgccg.nhs.uk/adults/investigations/inflammatory-marker-testing/*

Inflammatory markers are not useful in primary care as 'rule-out' tests.(1) For every 1000 inflammatory marker tests done there are 236 false positives, which leads to 710 GP appointments, 229 blood test appointments and 24 referrals in the following six months.

What do we recommend?

- Usually only one inflammatory marker test should be used.
- CRP should be the first line test in most circumstances (see table).
- Second line inflammatory markers differ across BNSSG; ESR is offered in Weston and UHBristol, plasma viscosity (PV) is offered by NBT. These should only be used for suspected temporal arteritis, as a second line test for polymyalgia rheumatica, for persistent bone pain in the over 60s or following secondary care request (see table below).
- Avoid using inflammatory markers for screening or as a rule-out for patients with non-specific symptoms.

Which test should be used?

Clinical question	CRP	ESR*	PV*	Comments
Screening asymptomatic patients	×	×	×	Unlikely to be useful. False positives common and may generate increased workload.
To 'rule-out' significant underlying disease in patients with non-specific symptoms e.g. tiredness	×	×	×	Inflammatory markers have low sensitivity and are therefore unsuitable as a rule-out test. False positives are common and may generate increased workload. (1) A small minority of patients with persistent tiredness will require referral to secondary care chronic fatigue/ME services and inflammatory markers are currently recommended prior to referral.
Could this patient have a significant infection?	✓	×	×	May be useful although not always necessary if symptoms and signs are clear cut. Point of care testing (if available) may reduce antibiotic prescribing in respiratory tract infections.(48)
Is this infection responding to antibiotic	*Little use*	×	×	For the vast majority of infections, repeat CRP testing is not indicated and assessment should be made on

Clinical question	CRP	ESR*	PV*	Comments
treatment?				clinical grounds. Monitoring of CRP may be useful is some chronic infections (e.g. osteomyelitis (4))
Does this patient have polymyalgia?	✓	*Second line*	*Second line*	CRP recommended first line. If CRP is normal and symptoms highly suggestive ESR or PV* should be added as a second line test.
Does this patient have giant cell arteritis?	✓	✓	✓	Both CRP and ESR/PV* are warranted – due to risks of serious complications if diagnosis is delayed.
Does this patient have inflammatory bowel disease?	*Little use*	×	×	A normal inflammatory marker is not a rule-out test. Calprotectin should be used first line if inflammatory bowel disease is suspected.
What is the cause of this/these inflamed joints?	✓	×	×	CRP is useful for secondary care triage of rheumatology referrals. However, normal inflammatory markers are not a rule-out; if clinical suspicion of inflammatory arthritis refer regardless of CRP results.
What is the cause of this patients raised platelets?	✓	×	×	British Society for Haematology recommends peripheral blood smear, CRP and iron studies as first line tests to investigate thrombocytosis. Raised CRP suggests reactive thrombocytosis, due to underlying inflammatory or malignant cause.(5)
Could this patient aged >60 with persistent bone or back pain or unexplained fracture have underlying pathology?	×	✓	✓	NICE recommend FBC, calcium and PV or ESR, to screen for potential underlying pathology including myeloma.(6, 7) If a raised inflammatory marker is detected then request a myeloma screen (see below).
Does this patient have myeloma?	×	×	×	If you suspect myeloma please order serum protein electrophoresis and urinary free light chains (Bence Jones protein) or serum free light chains, plus FBC, U+E, creatinine and calcium.
Monitoring of polymyalgia rheumatica	✓	×	×	Inflammatory markers are useful when tapering steroids. CRP is generally more sensitive than ESR or PV. No need to routinely test both simultaneously.

Clinical question	CRP	ESR*	PV*	Comments
Monitoring DMARDs	×	×	×	Inflammatory markers are not part of shared care protocols for DMARDs. Ask patients at each DMARD review if a) if they have a specialist out-patient appointment before the next routine blood test b) if their symptoms have flared such that they are needing to contact their GP or specialist team. Only do a CRP if the answer is yes to either one.
Blood tests prior to rheumatology secondary care review	✓	×	×	CRPs are useful in monitoring disease activity, and are needed to allow the specialists to calculate the disease activity score. They are therefore needed before each specialist review, but are unnecessary to monitor safety of medication on routine reviews, whether or not the dose has recently been changed.
Is this ***symptomatic*** patient with inflammatory arthritis or inflammatory bowel disease experiencing a flare?	✓	×	×	Raised inflammatory markers may be useful in confirming a disease flare and guiding appropriate management.

(*Weston and UHBristol offer ESR as a second line test; NBT offers plasma viscosity (PV))

How do I interpret a raised inflammatory marker in primary care?

Interpretation should be relatively straightforward if there is a clear pretest hypothesis against which the test result can be evaluated. The difficulty lies in the interpretation of an 'incidental' abnormality, when no specific disease is suspected. A systems inquiry, focusing on infection, autoimmune conditions, and malignancy, plus examination of the patient should generally point towards specific investigations.(8) If no obvious source can be found the test should be repeated. Men over 50 and women over 60 with persistently raised inflammatory markers have a cancer risk which exceeds the 3% NICE threshold for urgent investigation.(9) However inflammatory markers have a low sensitivity at <50%, so they are *not* recommended as a test to rule in or out the possibility of cancer in those with non-specific symptoms.

References

1. Watson J, Whiting P, Banks J, Pyne Y, Hamilton W. Added value and cascade effects of inflammatory marker tests in UK primary care: a cohort study from the Clinical Practice Research Datalink. BJGP 2019; 69 (684): e470-e478. DOI: 10.3399/bjgp19X704321
2. Watson J, Salisbury C, Whiting P, Banks J, Hamilton W. Use of multiple inflammatory marker tests in primary care: using Clinical Practice Research Datalink to evaluate accuracy. BJGP. 2019; 69 (684): e462-e469. DOI: https://doi.org/10.3399/bjgp19X704309
3. Eccles S, Pincus C, Higgins B, Woodhead M, Group GD. Diagnosis and management of community and hospital acquired pneumonia in adults: summary of NICE guidance. BMJ. 2014;349:g6722.

4. Van Asten SA, Nichols A, La Fontaine J, Bhavan K, Peters EJ, Lavery LA. The value of inflammatory markers to diagnose and monitor diabetic foot osteomyelitis. Int Wound J. 2017 Feb;14(1):40–5.
5. Mathur Abhinav, Samaranayake Shehan, Storrar Neill PF, Vickers Mark A. Investigating thrombocytosis BMJ 2019; 366 :l4183
6. NICE Guideline. Suspected Cancer: recognition and referral [NG12]. 2015. Available at: https://www.nice.org.uk/guidance/ng12/chapter/Recommendations-organised-by-symptom-and-findings-of-primary-care-investigations#skeletal-symptoms
7. Koshiaris C, Van den Bruel A, Oke JL, Nicholson BD, Shephard E, Braddick M, Hamilton W. Early detection of multiple myeloma in primary care using blood tests: a case–control study in primary care. BJGP 2018; 68 (674): e586-e593. DOI: 10.3399/bjgp18X698357
8. Watson J, Round A, Hamilton W. Raised inflammatory markers BMJ 2012; 344 :e454
9. Watson J, Salisbury C, Banks J, Whiting P, Hamilton W. Predictive value of inflammatory markers for cancer diagnosis in primary care: a prospective cohort study using electronic health records. British journal of cancer. 2019;120(11):1045-51.

Appendix F: Popups implemented in BNSSG electronic test ordering systems

ESR GP popup

Please select an indication for the ESR. A CRP will be requested where these indications are not met. For further details see: <https://remedy.bnssgccg.nhs.uk/adults/investigations/inflammatory-marker-testing/>

GP ESR Indications Please select an indication from the list below: (drop down)	Temporal arteritis/giant cell arteritis
	Polymyalgia rheumatica (second line testing)
	Persistent bone pain/unexplained fracture in >60yo
	Requirement from secondary care
	None of the above. A CRP will be requested

Plasma viscosity GP popup

Please select an indication for the plasma viscosity. A CRP will be requested where these indications are not met. For further details see: <https://remedy.bnssgccg.nhs.uk/adults/investigations/inflammatory-marker-testing/>

GP plasma viscosity Indications Please select an indication from the list below: (drop down)	Temporal arteritis/giant cell arteritis
	Polymyalgia rheumatica (second line testing)
	Persistent bone pain/unexplained fracture in >60yo
	Requirement from secondary care
	None of the above. A CRP will be requested

GP CRP popup

CRP tests are not useful as a 'rule-out' for patients with non-specific symptoms. For further details see: <https://remedy.bnssgccg.nhs.uk/adults/investigations/inflammatory-marker-testing/>

References

1. Watson J, Round, A., Hamilton, W. Raised inflammatory markers. BMJ. 2012;344:e454.

2. Watson J, de Salis I, Hamilton W, Salisbury C. 'I'm fishing really' - inflammatory marker testing in primary care: a qualitative study. BJGP. 2016;66(644):e200-6.

3. Syed M, Saleem T. Erythrocyte sedimentation rate: diagnostic value and pitfalls of usage in clinical practice. J Pak Med Assoc. 2010;60(2):156.

4. Tomlin A, Dovey S, Gauld R, Tilyard M. Better use of primary care laboratory services following interventions to 'market' clinical guidelines in New Zealand: a controlled before-and-after study. BMJ Quality & Safety. 2011;20(3):282-90.

5. Kesmarky G, Kenyeres P, Rabai M, Toth K. Plasma viscosity: a forgotten variable. Clin Hemorheol Microcirc. 2008;39(1-4):243-6.

6. Adelstein S, Baker, A. Making sense of inflammatory markers. Common Sense Pathology. 2014.

7. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. The Journal of Clinical Investigation. 2003;111(12):1805-12.

8. Hutchinson WL, Koenig W, Frohlich M, Sund M, Lowe GD, Pepys MBJCc. Immunoradiometric assay of circulating C-reactive protein: age-related values in the adult general population. 2000;46(7):934-8.

9. Yudkin JS, Stehouwer CDA, Emeis JJ, Coppack SW. C-Reactive Protein in Healthy Subjects: Associations With Obesity, Insulin Resistance, and Endothelial Dysfunction. Arteriosclerosis, Thombosis and Vascular Biology. 1999;19(4):972-8.

10. O'Sullivan JW, Stevens, S., Hobbs, F.D.R., Salisbury, C., Little, P., Goldacre, B., Bankhead, C., Aronson, J.K., Perera, R., Heneghan, C. Temporal trends in use of tests in UK primary care, 2000-15: retrospective analysis of 250 million tests. BMJ. 2018;363.

11. Busby J, Schroeder K, Woltersdorf W, Sterne JA, Ben-Shlomo Y, Hay A, et al. Temporal growth and geographic variation in the use of laboratory tests by NHS general practices: using routine data to identify research priorities. BJGP. 2013;63(609):e256-66.

12. O'Sullivan JW, Stevens S, Oke J, Hobbs FDR, Salisbury C, Little P, et al. Practice variation in the use of tests in UK primary care: a retrospective analysis of 16 million tests performed over 3.3 million patient years in 2015/16. 2018;16(1):229.

13. O'Sullivan JW, Heneghan C, Perera R, Oke J, Aronson JK, Shine B, et al. Variation in diagnostic test requests and outcomes: a preliminary metric for OpenPathology.net. Scientific reports. 2018;8(1):4752.

14. Salinas M, Lopez-Garrigos M, Diaz J, Ortuno M, Yago M, Laiz B, et al. Regional variations in test requiring patterns of general practitioners in Spain. Upsala journal of medical sciences. 2011;116(4):247-51.

15. Salinas M, Lopez-Garrigos M, Flores E, Uris J, Leiva-Salinas C. Request of acute phase markers in primary care in Spain. The American journal of managed care. 2015;21(10):e591-6.

16. Engstrom S, Molstad S, Lindstrom K, Nilsson G, Borgquist L. Excessive use of rapid tests in respiratory tract infections in Swedish primary health care. Scandinavian journal of infectious diseases. 2004;36(3):213-8.

17. Glasziou P, Straus S, Brownlee S, Trevena L, Dans L, Guyatt G, et al. Evidence for underuse of effective medical services around the world. The Lancet. 2017;390(10090):169-77.

18. NHS England. The 2nd Atlas of Variation in NHS Diagnostic Services in England. 2017.

19. Public Health England. NHS Atlas of Variation in Diagnostic Services. 2013.

20. Smellie WS, Galloway MJ, Chinn D, Gedling P. Is clinical practice variability the major reason for differences in pathology requesting patterns in general practice? J Clin Pathol. 2002;55(4):312-4.

21. Grytten J, Sorensen R. Practice variation and physician-specific effects. Journal of health economics. 2003;22(3):403-18.

22. Watson J, Bailey, S., Hamilton, F., Hamilton, W., Mounce L. Lessons from biases in electronic health record data: the importance of clinical vigilance with negative test results [letter]. BMJ. 2018;361.

23. O'Sullivan JW, Albasri A, Nicholson BD, Perera R, Aronson JK, Roberts N, et al. Overtesting and undertesting in primary care: a systematic review and meta-analysis. BMJ open. 2018;8(2):e018557.

24. Schattner A. The unbearable lightness of diagnostic testing: time to contain inappropriate test ordering. Postgraduate Medical Journal. 2008;84(998):618-21.

25. Deyo RA. Cascade effects of medical technology. Annual review of public health. 2002;23:23-44.

26. Naugler C, Ma, I. More than half of abnormal results from laboratory tests ordered by family physicians could be false-positive. J Canadian Family Physician. 2018;64(3):202-3.

27. Houben PH, van der Weijden T, Winkens RA, Grol RP. Cascade effects of laboratory testing are found to be rare in low disease probability situations: prospective cohort study. Journal of clinical epidemiology. 2010;63(4):452-8.

28. Spence D. Reducing general practice workload. BJGP. 2016;66(649):431.

29. Hobbs FDR, Bankhead C, Mukhtar T, Stevens S, Perera-Salazar R, Holt T, et al. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007-2014. The Lancet. 2016;387(10035):2323-30.

30. McAlister S, Barratt AL, Bell KJ, McGain F. The carbon footprint of pathology testing. The Medical journal of Australia. 2020;212(8):377-82.

31. Academy of Medical Royal Colleges. Protecting resources, promoting value: a doctor's guide to cutting waste in clinical care. 2014.

32. Salas RN, Maibach E, Pencheon D, Watts N, Frumkin H. A pathway to net zero emissions for healthcare. BMJ. 2020;371:m3785.

33. Gronlie M, Hjortdahl P. The erythrocyte sedimentation rate; its use and usefulness in primary health care. Scand J Prim Health Care. 1991;9(2):97-102.

34. Dahler-Eriksen BS, Lassen JF, Lund ED, Lauritzen T, Brandslund I. C-reactive protein in general practice--how commonly is it used and why? Scand J Prim Health Care. 1997;15(1):35-8.

35. Falk G, Fahey T. C-reactive protein and community-acquired pneumonia in ambulatory care: systematic review of diagnostic accuracy studies. Family Practice. 2009;26(1):10-21.

36. Ebell MH, McKay B, Guilbault R, Ermias Y. Diagnosis of acute rhinosinusitis in primary care: a systematic review of test accuracy. BJGP. 2016;66(650):e612-e32.

37. Kool M, Elshout G, Koes BW, Bohnen AM, Berger MY. C-Reactive Protein Level as Diagnostic Marker in Young Febrile Children Presenting in a General Practice Out-of-Hours Service. Journal of the American Board of Family Medicine : JABFM. 2016;29(4):460-8.

38. Hurst JR, Donaldson GC, Perera WR, Wilkinson TM, Bilello JA, Hagan GW, et al. Use of plasma biomarkers at exacerbation of chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine. 2006;174(8):867-74.

39. Craig JC, Williams GJ, Jones M, Codarini M, Macaskill P, Hayen A, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15781 febrile illnesses. BMJ. 2010;340.

40. van Vugt SF, Broekhuizen BD, Lammens C, Zuithoff NP, de Jong PA, Coenen S, et al. Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. BMJ. 2013;346:f2450.

41. Minnaard MC, de Groot JAH, Hopstaken RM, Schierenberg A, de Wit NJ, Reitsma JB, et al. The added value of C-reactive protein measurement in diagnosing pneumonia in primary care: a meta-analysis of individual patient data. CMAJ : Canadian Medical Association Journal. 2017;189(2):E56-e63.

42. Dinant GJ, Buntinx, F.F., Butler, C.C. The necessary shift from diagnostic to prognostic research. BMC Family Practice. 2007;8:53-.

43. Bruyndonckx R, Hens N, Verheij TJ, Aerts M, Ieven M, Butler CC, et al. Development of a prediction tool for patients presenting with acute cough in primary care: a prognostic study spanning six European countries. BJGP. 2018;68(670):e342-e50.

44. Aabenhus R, Jensen JU, Jorgensen KJ, Hrobjartsson A, Bjerrum L. Biomarkers as point-of-care tests to guide prescription of antibiotics in patients with acute respiratory infections in primary care. The Cochrane database of systematic reviews. 2014;11:Cd010130.

45. Butler CC, Gillespie D, White P, Bates J, Lowe R, Thomas-Jones E, et al. C-Reactive Protein Testing to Guide Antibiotic Prescribing for COPD Exacerbations. NEJM. 2019;381(2):111-20.

46. Little P, Stuart B, Francis N, Douglas E, Tonkin-Crine S, Anthierens S, et al. Antibiotic Prescribing for Acute Respiratory Tract Infections 12 Months After Communication and CRP Training: A Randomized Trial. Ann Fam Med. 2019;17(2):125-32.

47. Huddy JR, Ni MZ, Barlow J, Majeed A, Hanna GB. Point-of-care C reactive protein for the diagnosis of lower respiratory tract infection in NHS primary care: a qualitative study of barriers and facilitators to adoption. BMJ open. 2016;6(3).

48. Eccles S, Pincus C, Higgins B, Woodhead M, Group GD. Diagnosis and management of community and hospital acquired pneumonia in adults: summary of NICE guidance. BMJ. 2014;349:g6722.

49. Rebnord IK, Hunskaar S, Gjesdal S, Hetlevik O. Point-of-care testing with CRP in primary care: a registry-based observational study from Norway. BMC Fam Pract. 2015;16(1):170.

50. Buttgereit F, Dejaco C, Matteson EL, Dasgupta B. Polymyalgia Rheumatica and Giant Cell Arteritis: A Systematic Review. Jama. 2016;315(22):2442-58.

51. Hassan N, Dasgupta, B., Barraclough, K. Giant cell arteritis. BMJ. 2011;342:d3019.

52. Helliwell T, Muller S, Hider S. ESR can be normal in giant cell arteritis and polymyalgia rheumatica. BMJ. 2012;344:e1408; author reply e9.

53. Parikh M, Miller NR, Lee AG, Savino PJ, Vacarezza MN, Cornblath W, et al. Prevalence of a normal C-reactive protein with an elevated erythrocyte sedimentation rate in biopsy-proven giant cell arteritis. Ophthalmology. 2006;113(10):1842-5.

54. Dasgupta B. Concise guidance: diagnosis and management of giant cell arteritis. J Clinical Medicine. 2010;10(4):381-6.

55. Deighton C, O'Mahony R, Tosh J, Turner C, Rudolf M, Group GD. Management of rheumatoid arthritis: summary of NICE guidance. BMJ. 2009;338:b702.

56. Sokka T, Pincus T. Erythrocyte Sedimentation Rate, C-Reactive Protein, or Rheumatoid Factor Are Normal at Presentation in 35%-45% of Patients with Rheumatoid Arthritis Seen Between 1980 and 2004: Analyses from Finland and the United States. J Rheumatol. 2009;36(7):1387-90.

57. National Institute for Health and Care Excellence. Faecal calprotectin diagnostic tests for inflammatory diseases of the bowel. Diagnostics guidance [DG11]. 2013.

58. Watson J, Mounce L, Bailey SE, Cooper SL, Hamilton W. Blood markers for cancer. BMJ. 2019;367:l5774.

59. Koshiaris C, Van den Bruel A, Oke JL, Nicholson BD, Shephard E, Braddick M, et al. Early detection of multiple myeloma in primary care using blood tests: a case–control study in primary care. BJGP. 2018;68(674):e586-e93.

60. Stapley S, Peters T, Neal R, Rose P, Walter F, Hamilton W. The risk of pancreatic cancer in symptomatic patients in primary care: a large case–control study using electronic records. British journal of cancer. 2012;106(12):1940-4.

61. Shephard EA, Stapley S, Neal RD, Rose P, Walter FM, Hamilton WT. Clinical features of bladder cancer in primary care. BJGP. 2012;62(602):e598-e604.

62. Shephard E, Neal R, Rose P, Walter F, Hamilton WT. Clinical features of kidney cancer in primary care: a case-control study using primary care records. Br J Gen Pract. 2013;63(609):e250-e5.

63. Shephard EA, Neal RD, Rose PW, Walter FM, Hamilton WT. Quantifying the risk of non-Hodgkin lymphoma in symptomatic primary care patients aged ⩾40 years: a large case‐control study using electronic records. BJGP. 2015;65(634):e281-e8.

64. Guo YZ, Pan L, Du CJ, Ren DQ, Xie XM. Association between C-reactive protein and risk of cancer: a meta-analysis of prospective cohort studies. Asian Pacific journal of cancer prevention : APJCP. 2013;14(1):243-8.

65. National Institute for Health and Clinical Excellence. Suspected cancer: recognition and referral [NG12]. 2015.

66. van der Meer V, Neven AK, van den Broek PJ, Assendelft WJJ. Diagnostic value of C reactive protein in infections of the lower respiratory tract: systematic review. BMJ. 2005;331(7507):26.

67. Engel MF, Paling FP, Hoepelman AI, van der Meer V, Oosterheert JJ. Evaluating the evidence for the implementation of C-reactive protein measurement in adult patients with suspected lower respiratory tract infection in primary care: a systematic review. Fam Pract. 2012;29(4):383-93.

68. Flood RG, Badik, J., Aronoff, S.C. The utility of serum C-reactive protein in differentiating bacterial from nonbacterial pneumonia in children: a meta-analysis of 1230 children. Pediatr Infect Dis J. 2008;27(2):95-9.

69. Chen YW, Leung JM, Sin DD. A Systematic Review of Diagnostic Biomarkers of COPD Exacerbation. PloS one. 2016;11(7):e0158843.

70. Shoki AH, Mayer-Hamblett N, Wilcox PG, Sin DD, Quon BS. Systematic review of blood biomarkers in cystic fibrosis pulmonary exacerbations. Chest. 2013;144(5):1659-70.

71. Yoon C, Chaisson LH, Patel SM, Allen IE, Drain PK, Wilson D, et al. Diagnostic accuracy of C-reactive protein for active pulmonary tuberculosis: a meta-analysis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2017;21(9):1013-9.

72. Bundy DG, Byerley JS, Liles EA, Perrin EM, Katznelson J, Rice HE. Does this child have appendicitis? Jama. 2007;298(4):438-51.

73. Yu CW, Juan LI, Wu MH, Shen CJ, Wu JY, Lee CC. Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. The British journal of surgery. 2013;100(3):322-9.

74. Gans SL, Atema JJ, Stoker J, Toorenvliet BR, Laurell H, Boermeester MA. C-reactive protein and white blood cell count as triage test between urgent and nonurgent conditions in 2961 patients with acute abdominal pain. Medicine. 2015;94(9):e569.

75. Cousin F, Ortega-Deballon P, Bourredjem A, Doussot A, Giaccaglia V, Fournel I. Diagnostic Accuracy of Procalcitonin and C-reactive Protein for the

Early Diagnosis of Intra-abdominal Infection After Elective Colorectal Surgery: A Meta-analysis. Annals of surgery. 2016;264(2):252-6.

76. Gans SL, Atema JJ, van Dieren S, Groot Koerkamp B, Boermeester MA. Diagnostic value of C-reactive protein to rule out infectious complications after major abdominal surgery: a systematic review and meta-analysis. International journal of colorectal disease. 2015;30(7):861-73.

77. Yang SK, Xiao L, Zhang H, Xu XX, Song PA, Liu FY, et al. Significance of serum procalcitonin as biomarker for detection of bacterial peritonitis: a systematic review and meta-analysis. BMC infectious diseases. 2014;14:452.

78. Lin KH, Wang FL, Wu MS, Jiang BY, Kao WL, Chao HY, et al. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection in patients with liver cirrhosis: a systematic review and meta-analysis. Diagnostic microbiology and infectious disease. 2014;80(1):72-8.

79. Adamina M, Steffen T, Tarantino I, Beutner U, Schmied BM, Warschkow R. Meta-analysis of the predictive value of C-reactive protein for infectious complications in abdominal surgery. The British journal of surgery. 2015;102(6):590-8.

80. Straatman J, Harmsen AM, Cuesta MA, Berkhof J, Jansma EP, van der Peet DL. Predictive Value of C-Reactive Protein for Major Complications after Major Abdominal Surgery: A Systematic Review and Pooled-Analysis. PloS one. 2015;10(7):e0132995.

81. Singh PP, Zeng IS, Srinivasa S, Lemanu DP, Connolly AB, Hill AG. Systematic review and meta-analysis of use of serum C-reactive protein levels to predict anastomotic leak after colorectal surgery. The British journal of surgery. 2014;101(4):339-46.

82. Tan M, Lu Y, Jiang H, Zhang L. The diagnostic accuracy of procalcitonin and C-reactive protein for sepsis: A systematic review and meta-analysis. Journal of cellular biochemistry. 2019;120(4):5852-9.

83. Ruan L, Chen G-Y, Liu Z, Zhao Y, Xu G-Y, Li S-F, et al. The combination of procalcitonin and C-reactive protein or presepsin alone improves the accuracy of diagnosis of neonatal sepsis: a meta-analysis and systematic review. Critical Care. 2018;22(1):316.

84. Brown JVE, Meader N, Cleminson J, McGuire W. C-reactive protein for diagnosing late-onset infection in newborn infants. The Cochrane database of systematic reviews. 2019;1(1):Cd012126.

85. Henry BM, Roy J, Ramakrishnan PK, Vikse J, Tomaszewski KA, Walocha JA. Procalcitonin as a Serum Biomarker for Differentiation of Bacterial Meningitis

From Viral Meningitis in Children: Evidence From a Meta-Analysis. Clin Pediatr (Phila). 2016;55(8):749-64.

86. Vikse J, Henry BM, Roy J, Ramakrishnan PK, Tomaszewski KA, Walocha JA. The role of serum procalcitonin in the diagnosis of bacterial meningitis in adults: a systematic review and meta-analysis. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2015;38:68-76.

87. Wu CW, Wu JY, Chen CK, Huang SL, Hsu SC, Lee MG, et al. Does procalcitonin, C-reactive protein, or interleukin-6 test have a role in the diagnosis of severe infection in patients with febrile neutropenia? A systematic review and meta-analysis. Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer. 2015.

88. Phillips RS, Wade R, Lehrnbecher T, Stewart LA, Sutton AJ. Systematic review and meta-analysis of the value of initial biomarkers in predicting adverse outcome in febrile neutropenic episodes in children and young people with cancer. BMC medicine. 2012;10:6.

89. Haeusler GM, Carlesse F, Phillips RS. An updated systematic review and meta-analysis of the predictive value of serum biomarkers in the assessment of fever during neutropenia in children with cancer. Pediatr Infect Dis J. 2013;32(10):e390-6.

90. Shaikh KJ, Osio VA, Leeflang MM, Shaikh N. Procalcitonin, C-reactive protein, and erythrocyte sedimentation rate for the diagnosis of acute pyelonephritis in children. The Cochrane database of systematic reviews. 2020;9:Cd009185.

91. Van den Bruel A, Thompson MJ, Haj-Hassan T, Stevens R, Moll H, Lakhanpaul M, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. BMJ. 2011;342.

92. Sanders S, Barnett A, Correa-Velez I, Coulthard M, Doust J. Systematic review of the diagnostic accuracy of C-reactive protein to detect bacterial infection in nonhospitalized infants and children with fever. J Pediatr. 2008;153(4):570-4.

93. Yo CH, Hsieh PS, Lee SH, Wu JY, Chang SS, Tasi KC, et al. Comparison of the test characteristics of procalcitonin to C-reactive protein and leukocytosis for the detection of serious bacterial infections in children presenting with fever without source: a systematic review and meta-analysis. Annals of emergency medicine. 2012;60(5):591-600.

94. Shen CJ, Wu MS, Lin KH, Lin WL, Chen HC, Wu JY, et al. The use of procalcitonin in the diagnosis of bone and joint infection: a systemic review and meta-analysis. European journal of clinical microbiology & infectious diseases :

official publication of the European Society of Clinical Microbiology. 2013;32(6):807-14.

95. Carpenter CR, Schuur JD, Everett WW, Pines JM. Evidence-based diagnostics: adult septic arthritis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine. 2011;18(8):781-96.

96. Zhao J, Zhang S, Zhang L, Dong X, Li J, Wang Y, et al. Serum procalcitonin levels as a diagnostic marker for septic arthritis: A meta-analysis. The American journal of emergency medicine. 2017;35(8):1166-71.

97. Victoria van Asten SA, Geradus Peters EJ, Xi Y, Lavery LA. The Role of Biomarkers to Diagnose Diabetic Foot Osteomyelitis. A Meta-analysis. Current diabetes reviews. 2016;12(4):396-402.

98. Yuan K, Chen HL, Cui ZM. Diagnostic accuracy of C-reactive protein for periprosthetic joint infection: a meta-analysis. Surgical infections. 2014;15(5):548-59.

99. Berbari E, Mabry T, Tsaras G, Spangehl M, Erwin PJ, Murad MH, et al. Inflammatory blood laboratory levels as markers of prosthetic joint infection: a systematic review and meta-analysis. J Bone Joint Surg Am. 2010;92(11):2102-9.

100. Etyang AK, Omuse G, Mukaindo AM, Temmerman M. Maternal inflammatory markers for chorioamnionitis in preterm prelabour rupture of membranes: a systematic review and meta-analysis of diagnostic test accuracy studies. Systematic reviews. 2020;9(1):141.

101. Lu XL, Xiao ZH, Yang MY, Zhu YM. Diagnostic value of serum procalcitonin in patients with chronic renal insufficiency: a systematic review and meta-analysis. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2013;28(1):122-9.

102. Lyu YX, Yu XC, Zhu MY. Comparison of the diagnostic value of procalcitonin and C-reactive protein after hematopoietic stem cell transplantation: a systematic review and meta-analysis. Transplant Infectious Disease. 2013;15(3):290-9.

103. Yu CW, Juan LI, Hsu SC, Chen CK, Wu CW, Lee CC, et al. Role of procalcitonin in the diagnosis of infective endocarditis: a meta-analysis. The American journal of emergency medicine. 2013;31(6):935-41.

104. Menees SB, Powell C, Kurlander J, Goel A, Chey WD. A meta-analysis of the utility of C-reactive protein, erythrocyte sedimentation rate, fecal calprotectin, and fecal lactoferrin to exclude inflammatory bowel disease in adults with IBS. Am J Gastroenterol. 2015;110(3):444-54.

105. Jellema P, van Tulder MW, van der Horst HE, Florie J, Mulder CJ, van der Windt DAWM. Inflammatory bowel disease: a systematic review on the value of diagnostic testing in primary care. Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland. 2011;13(3):239-54.

106. Holtman GA, Lisman-van Leeuwen Y, Reitsma JB, Berger MY. Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis. Pediatrics. 2016;137(1).

107. Song GG, Bae, S.C., & Lee, Y.H. Diagnostic accuracies of procalcitonin and C-reactive protein for bacterial infection in patients with systemic rheumatic diseases: a meta-analysis. Clinical and experimental rheumatology. 2015;33(2):166-73.

108. Stegeman I, Ochodo EA, Guleid F, Holtman GA, Yang B, Davenport C, et al. Routine laboratory testing to determine if a patient has COVID-19. The Cochrane database of systematic reviews. 2020;11:Cd013787.

109. Allen A, Carville S, McKenna FJ. Diagnosis and management of rheumatoid arthritis in adults: summary of updated NICE guidance. BMJ. 2018;362:k3015.

110. van Asten SA, Jupiter DC, Mithani M, La Fontaine J, Davis KE, Lavery LA. Erythrocyte sedimentation rate and C-reactive protein to monitor treatment outcomes in diabetic foot osteomyelitis. International Wound Journal. 2017;14(1):142-8.

111. Allin KH, Nordestgaard BG. Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer. Critical reviews in clinical laboratory sciences. 2011;48(4):155-70.

112. Shrotriya S, Walsh D, Bennani-Baiti N, Thomas S, Lorton C. C-Reactive Protein Is an Important Biomarker for Prognosis Tumor Recurrence and Treatment Response in Adult Solid Tumors: A Systematic Review. PloS one. 2015;10(12):e0143080.

113. GP Update. Hypertension and ambulatory BPs. 2011 [Available from: https://www.gp-update.co.uk/Latest-Updates/Hypertension-and-ambulatory-BPs].

114. NICE Clinical Knowledge Summaries. Dementia. 2020 [Available from: https://cks.nice.org.uk/dementia]

115. NICE Clinical Knowledge Summaries. Tiredness/fatigue in adults. 2020 [Available from: https://cks.nice.org.uk/tirednessfatigue-in-adults].

116. National Institute for Health and Care Excellence. Irritable bowel syndrome in adults: diagnosis and management [CG61] 2008 (Updated 2017) [Available from: https://www.nice.org.uk/guidance/cg61]

117. Koch H, van Bokhoven MA, ter Riet G, van Alphen-Jager JT, van der Weijden T, Dinant GJ, et al. Ordering blood tests for patients with unexplained fatigue in general practice: what does it yield? Results of the VAMPIRE trial. Br J Gen Pract. 2009;59(561):e93-100.

118. Deyo RA. Cascade effects of medical technology. Annual review of public health. 2002;23:23-44.

119. Froom P, Margaliot S, Caine Y, Benbassat J. Significance of erythrocyte sedimentation rate in young adults. Am J Clin Pathol. 1984;82(2):198-200.

120. Thomas PD, Goodwin JS. Diagnostic importance of an elevated erythrocyte sedimentation rate in the elderly. Clin Rheumatol. 1987;6(2):177-80.

121. Dinant GJ, Knottnerus JA, Van Wersch JW. Discriminating ability of the erythrocyte sedimentation rate: a prospective study in general practice. BJGP. 1991;41(350):365-70.

122. Dinant GJ, Knottnerus JA, Van Wersch JW. Diagnostic impact of the erythrocyte sedimentation rate in general practice: a before-after analysis. Family Practice. 1992;9(1):28-31.

123. Kirkeby OJ, Risoe C, Vikland R. Significance of a high erythrocyte sedimentation rate in general practice. Br J Clin Pract. 1989;43(7):252-4.

124. Katz PR, Gutman SI, Richman G, Karuza J, Bartholomew WR, Baum J. Erythrocyte sedimentation rate and C-reactive protein compared in the elderly. Clin Chem. 1989;35(3):466-8.

125. Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, Collins R, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. Lancet. 2010;375(9709):132-40.

126. Wang X, Bao W, Liu J, Ouyang YY, Wang D, Rong S, et al. Inflammatory markers and risk of type 2 diabetes: a systematic review and meta-analysis. Diabetes Care. 2013;36(1):166-75.

127. Wu N, Xu B, Xiang Y, Wu L, Zhang Y, Ma X, et al. Association of inflammatory factors with occurrence and recurrence of atrial fibrillation: a meta-analysis. International journal of cardiology. 2013;169(1):62-72.

128. Zhou B, Liu J, Wang ZM, Xi T. C-reactive protein, interleukin 6 and lung cancer risk: a meta-analysis. PloS one. 2012;7(8):e43075.

129. Zhou B, Shu B, Yang J, Liu J, Xi T, Xing Y. C-reactive protein, interleukin-6 and the risk of colorectal cancer: a meta-analysis. Cancer causes & control : CCC. 2014;25(10):1397-405.

130. Guo L, Liu S, Zhang S, Chen Q, Zhang M, Quan P, et al. C-reactive protein and risk of breast cancer: A systematic review and meta-analysis. Scientific reports. 2015;5:10508.

131. Koyama A, O'Brien J, Weuve J, Blacker D, Metti AL, Yaffe K. The role of peripheral inflammatory markers in dementia and Alzheimer's disease: a meta-analysis. The journals of gerontology Series A, Biological sciences and medical sciences. 2013;68(4):433-40.

132. Metcalf SA, Jones PB, Nordstrom T, Timonen M, Maki P, Miettunen J, et al. Serum C-reactive protein in adolescence and risk of schizophrenia in adulthood: A prospective birth cohort study. Brain, behavior, and immunity. 2017;59:253-9.

133. Valkanova V, Ebmeier KP, Allan CL. CRP, IL-6 and depression: a systematic review and meta-analysis of longitudinal studies. Journal of affective disorders. 2013;150(3):736-44.

134. Li Y, Zhong X, Cheng G, Zhao C, Zhang L, Hong Y, et al. Hs-CRP and all-cause, cardiovascular, and cancer mortality risk: a meta-analysis. Atherosclerosis. 2017;259:75-82.

135. Hippisley-Cox J, Coupland C. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. BMJ open. 2013;3(8):e003482.

136. Hippisley-Cox J, Coupland C. Development and validation of QMortality risk prediction algorithm to estimate short term risk of death and assess frailty: cohort study. BMJ. 2017;358:j4208.

137. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and Ageing. 2016;45(3):353-60.

138. Yoshinaga R, Doi Y, Ayukawa K, Ishikawa S. High-sensitivity C reactive protein as a predictor of inhospital mortality in patients with cardiovascular disease at an emergency department: a retrospective cohort study. BMJ open. 2017;7(10).

139. Barron E, Lara J, White M, Mathers JC. Blood-borne biomarkers of mortality risk: systematic review of cohort studies. PloS one. 2015;10(6):e0127550.

140. Genest J, McPherson R, Frohlich J, Anderson T, Campbell N, Carpentier A, et al. 2009 Canadian Cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult – 2009 recommendations. Canadian Journal of Cardiology. 2009;25(10):567-79.

141. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. Circulation. 2013:01. cir. 0000437741.48606. 98.

142. Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, Criqui M, et al. Markers of inflammation and cardiovascular disease. Circulation. 2003;107(3):499-511.

143. Lapić I, Padoan A, Bozzato D, Plebani M. Erythrocyte Sedimentation Rate and C-Reactive Protein in Acute Inflammation: Meta-Analysis of Diagnostic Accuracy Studies. Am J Clin Pathol. 2019;153(1):14-29.

144. Assasi N, Blackhouse G, Campbell K, Hopkins RB, Levine M, Richter T, et al. Comparative Value of Erythrocyte Sedimentation Rate (ESR) and C-Reactive Protein (CRP) Testing in Combination Versus Individually for the Diagnosis of Undifferentiated Patients With Suspected Inflammatory Disease or Serious Infection: A Systematic Review and Economic Analysis. CADTH Health Technology Assessments. 2015.

145. Bajada S, Yoong AWH, Hourigan P, Koopmans PC, Phillips JRA, Toms AD. Plasma Viscosity Has a Role in the Diagnosis of Prosthetic Joint Infection After Total Knee Arthroplasty. The Journal of arthroplasty. 2019;34(12):3035-9.

146. Vercoutere W, Thevissen K, Bombardier C, Landewé RB. Diagnostic and predictive value of acute-phase reactants in adult undifferentiated peripheral inflammatory arthritis: a systematic review. The Journal of rheumatology Supplement. 2011;87:15-9.

147. Ford MJ, Innes JA, Parrish FM, Allan NC, Horn DB, Munro JF. The significance of gross elevations of the erythrocyte sedimentation rate in a general medical unit. Eur J Clin Invest. 1979;9(3):191-4.

148. Monig H, Marquardt D, Arendt T, Kloehn S. Limited value of elevated erythrocyte sedimentation rate as an indicator of malignancy. Family Practice. 2002;19(5):436-8.

149. Fincher RM, Page MI. Clinical significance of extreme elevation of the erythrocyte sedimentation rate. Arch Intern Med. 1986;146(8):1581-3.

150. Lluberas-Acosta G, Schumacher HR, Jr. Markedly elevated erythrocyte sedimentation rates: consideration of clinical implications in a hospital population. Br J Clin Pract. 1996;50(3):138-42.

151. Wyler DJ. Diagnostic implications of markedly elevated erythrocyte sedimentation rate: a reevaluation. South Med J. 1977;70(12):1428-30.

152. Landry A, Docherty P, Ouellette S, Cartier LJ. Causes and outcomes of markedly elevated C-reactive protein levels. Canadian family physician Medecin de famille canadien. 2017;63(6):e316-e23.

153. Little P, Cantrell T, Roberts L, Chapman J, Langridge J, Pickering R. Why do GPs perform investigations?: The medical and social agendas in arranging back X-rays. Fam Pract. 1998;15(3):264-5.

154. Fryer AA, Smellie WSA. Managing demand for laboratory tests: a laboratory toolkit. J Clin Pathol. 2013;66(1):62-72.

155. Callen JL, Westbrook JI, Georgiou A, Li J. Failure to follow-up test results for ambulatory patients: a systematic review. Journal of general internal medicine. 2012;27(10):1334-48.

156. Litchfield IJ, Bentham LM, Lilford RJ, Greenfield SM. Test result communication in primary care: clinical and office staff perspectives. Fam Pract. 2014;31(5):592-7.

157. Litchfield I, Bentham L, Hill A, McManus RJ, Lilford R, Greenfield S. Routine failures in the process for blood testing and the communication of results to patients in primary care in the UK: a qualitative exploration of patient and provider perspectives. BMJ Quality & Safety. 2015;24(11):681-90.

158. Litchfield I, Bentham L, Lilford R, McManus RJ, Hill A, Greenfield S. Test result communication in primary care: a survey of current practice. BMJ Quality & Safety. 2015;24(11):691-9.

159. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care--a systematic review. Fam Pract. 2008;25(6):400-13.

160. Barry MJ, Edgman-Levitan S. Shared Decision Making — The Pinnacle of Patient-Centered Care. New England Journal of Medicine. 2012;366(9):780-1.

161. Légaré F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database of Systematic Reviews. 2018(7).

162. Whiting P, Toerien M, de Salis I, Sterne JA, Dieppe P, Egger M, et al. A review identifies and classifies reasons for ordering diagnostic tests. Journal of clinical epidemiology. 2007;60(10):981-9.

163. Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: practicing physicians' use of quantitative measures of test accuracy. Am J Med. 1998;104(4):374-80.

164. Watson J, de Salis I, Banks J, Salisbury C. What do tests do for doctors? A qualitative study of blood testing in UK primary care. Family Practice. 2017;34(6):735-9.

165. Duddy C, Wong G. Efficiency over thoroughness in laboratory testing decision-making in primary care: findings from a realist review. BJGP Open. 2021:doi: 10.3399/bjgpopen20X101146.

166. van Bokhoven MA, Pleunis-van Empel MC, Koch H, Grol RP, Dinant GJ, van der Weijden T. Why do patients want to have their blood tested? A qualitative study of patient expectations in general practice. BMC Fam Pract. 2006;7:75.

167. Hoffmann TC, Del Mar C. Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA internal medicine. 2015;175(2):274-86.

168. Howard L, Wessely S, Leese M, Page L, McCrone P, Husain K, et al. Are investigations anxiolytic or anxiogenic? A randomised controlled trial of neuroimaging to provide reassurance in chronic daily headache. Journal of Neurology, Neurosurgery, and Psychiatry. 2005;76(11):1558-64.

169. Rolfe A, Burton C. Reassurance after diagnostic testing with a low pretest probability of serious disease: Systematic review and meta-analysis. JAMA internal medicine. 2013;173(6):407-16.

170. Petrie KJ, Sherriff R. Normal diagnostic test results do not reassure patients. BMJ Evidence Based Medicine. 2013(19):14.

171. Cabral C, Horwood J, Hay AD, Lucas PJ. How communication affects prescription decisions in consultations for acute illness in children: a systematic review and meta-ethnography. BMC Fam Pract. 2014;15:63.

172. Williams S, Weinman J, Dale J, Newman S. Patient expectations: what do primary care patients want from the GP and how far does meeting expectations affect patient satisfaction? Fam Pract. 1995;12(2):193-201.

173. General Medical Council. Decision making and consent. 2020 [Available from: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/decision-making-and-consent].

174. United Kingdom Supreme Court. Montgomery (Appellant) v Lanarkshire Health Board (Respondent) (Scotland) 2015 [Available from: http://www.bailii.org/uk/cases/UKSC/2015/11.html].

175. MDU. Consent to blood tests. 2002 [Available from: https://www.themdu.com/guidance-and-advice/case-studies/consent-to-blood-tests].

176. National Institute for Health and Care Excellence. Shared decision making [In development NG10120] 2021 [Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ng10120].

177. James LJ, Wong G, Craig JC, Hanson CS, Ju A, Howard K, et al. Men's perspectives of prostate cancer screening: A systematic review of qualitative studies. PloS one. 2017;12(11):e0188258.

178. Medendorp NM, van Maarschalkerweerd PEA, Murugesu L, Daams JG, Smets EMA, Hillen MA. The impact of communicating uncertain test results in cancer genetic counseling: A systematic mixed studies review. Patient Educ Couns. 2020;103(9):1692-708.

179. Edwards AGK, Naik G, Ahmed H, Elwyn GJ, Pickles T, Hood K, et al. Personalised risk communication for informed decision making about taking screening tests. Cochrane Database of Systematic Reviews. 2013(2).

180. Riikonen JM, Guyatt GH, Kilpeläinen TP, Craigie S, Agarwal A, Agoritsas T, et al. Decision Aids for Prostate Cancer Screening Choice: A Systematic Review and Meta-analysis. JAMA internal medicine. 2019;179(8):1072-82.

181. Martin JE, Watson J, Barnes RK. Shared decision making about blood tests: secondary analysis of video-recorded primary care consultations. Br J Gen Pract. 2020;70(694):e339-e47.

182. Ford S, Schofield T, Hope T. Observing decision-making in the general practice consultation: who makes which decisions? Health expectations : an international journal of public participation in health care and health policy. 2006;9(2):130-7.

183. Usher-Smith JA, Sharp, Stephen J, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. BMJ. 2016;353.

184. Whiting P, Davenport CJD, Research P. Understanding test accuracy research: a test consequence graphic. 2018;2(1):2.

185. Denzin NK, Lincoln YS. The Sage handbook of qualitative research: sage; 2011.

186. McEvoy P, Richards D. A critical realist rationale for using a combination of quantitative and qualitative methods. Journal of Research in Nursing. 2006;11(1):66-78.

187. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.

188. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. British journal of clinical pharmacology. 2010;69(1):4-14.

189. Bailey SE, Ukoumunne OC, Shephard EA, Hamilton W. Clinical relevance of thrombocytosis in primary care: a prospective cohort study of cancer incidence

using English electronic medical records and cancer registry data. Br J Gen Pract. 2017.

190. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Clin Chem. 2015;61(12)(1446-1452).

191. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885.

192. Watson J, Nicholson BD, Hamilton W, Price S. Identifying clinical features in primary care electronic health record studies: methods for codelist development. BMJ open. 2017;7(11).

193. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J. Consensus development methods, and their use in clinical guideline development. 1998;2(3).

194. Watson J. CPRD codes: infections and autoimmune disease. 2018 [Available from: https://doi.org/10.5523/bris.2954m5h0ync672u8yzx16xxj7l].

195. University of Exeter. Cancer diagnosis (DISCO) [Available from: http://medicine.exeter.ac.uk/research/healthresearch/disco/].

196. The CanTest Collaborative. CanTest 2021 [Available from: https://www.cantest.org/].

197. Mackay IR, Rose NR. The autoimmune diseases: Elsevier Inc.; 2013.

198. American Autoimmune Related Diseases Association. Autoimmune Disease List [Available from: https://www.aarda.org/diseaselist/].

199. StataCorp. Stata Statistical Software: Release 15. . College Station, TX: StataCorp LLC; 2017.

200. Seed P. DIAGT: Stata module to report summary statistics for diagnostic tests compared to true disease status. 2010.

201. Whiting P, Martin RM, Ben-Shlomo Y, Gunnell D, Sterne JAC. How to apply the results of a research paper on diagnosis to your patient. JRSM short reports. 2013;4(1):7.

202. DeLong ER, DeLong DM, Clarke-Pearson DLJB. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. 1988:837-45.

203. Office for National Statistics. Populations Estimates for UK, England and Wals, Scotland and Northern Ireland, Mid-2014. 2014.

204. Hippisley-Cox J. ONS Consultations Report 20 - QResearch calendar year consultations 2008. 2008.

205. Office for National Statistics. Cancer Registration Statistics, England: 2014 2014 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014#the-most-common-cancers-registered-were-breast-prostate-lung-and-colorectal-cancers].

206. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health. 2019:1-16.

207. Braun V, Clarke V. Using theatic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.

208. Elo S, Kyngäs H. The qualitative content analysis process. Journal of advanced nursing. 2008;62(1):107-15.

209. Tudor Hart J. THE INVERSE CARE LAW. The Lancet. 1971;297(7696):405-12.

210. Whyte MB, Hinton W, McGovern A, van Vlymen J, Ferreira F, Calderara S, et al. Disparities in glycaemic control, monitoring, and treatment of type 2 diabetes in England: A retrospective cohort analysis. PLoS Med. 2019;16(10):e1002942.

211. Smith D, Thomson K, Bambra C, Todd A. The breast cancer paradox: A systematic review of the association between area-level deprivation and breast cancer screening uptake in Europe. Cancer Epidemiology. 2019;60:77-85.

212. Morgan RM, Steele RJ, Nabi G, McCowan C. Socioeconomic variation and prostate specific antigen testing in the community: a United Kingdom based population study. The Journal of urology. 2013;190(4):1207-12.

213. Spence D. The inverse care law has had its day. BMJ. 2007;334(7592):538-.

214. Hamilton W. Earlier diagnosis of colorectal, lung and prostate cancer [MD thesis]: Bristol; 2005.

215. Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. BMJ. 2018;361.

216. Barth JH, Jones RG. Indiscriminate investigations have adverse effects. BMJ. 2003;326(7385):393.

217. Sah S, Elias P, Ariely D. Investigation momentum: the relentless pursuit to resolve uncertainty. JAMA internal medicine. 2013;173(10):932-3.

218. Heikkila K, Harris R, Lowe G, Rumley A, Yarnell J, Gallacher J, et al. Associations of circulating C-reactive protein and interleukin-6 with cancer risk: findings from two prospective cohorts and a meta-analysis. Cancer causes & control : CCC. 2009;20(1):15-26.

219. Zhang X, Liu S, Zhou Y. Circulating levels of C-reactive protein, interleukin-6 and tumor necrosis factor-alpha and risk of colorectal adenomas: a meta-analysis. Oncotarget. 2016;7(39):64371-9.

220. Zeng F, Wei H, Yeoh E, Zhang Z, Ren ZF, Colditz GA, et al. Inflammatory Markers of CRP, IL6, TNFalpha, and Soluble TNFR2 and the Risk of Ovarian Cancer: A Meta-analysis of Prospective Studies. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2016;25(8):1231-9.

221. Chan DS, Bandera EV, Greenwood DC, Norat T. Circulating C-reactive protein and breast cancer risk - systematic literature review and meta-analysis of prospective cohort studies. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2015.

222. Shephard EA, Neal RD, Rose P, Walter FM, Litt EJ, Hamilton WT. Quantifying the risk of multiple myeloma from symptoms reported in primary care patients: a large case–control study using electronic records. Br J Gen Pract. 2015;65(631):e106-e13.

223. Hartmann RLJ, Martin S, Anne G, Steen L, Peter G, Kasper I, et al. Inflammatory biomarkers and cancer: CRP and suPAR as markers of incident cancer in patients with serious nonspecific symptoms and signs of cancer. International Journal of Cancer. 2017;141(1):191-9.

224. Zacho J, Tybjaerg-Hansen A, Nordestgaard BG. C-reactive protein and all-cause mortality-the Copenhagen City Heart Study. Eur Heart J. 2010;31(13):1624-32.

225. Ford ES, Cunningham TJ, Mannino DM. Inflammatory markers and mortality among US adults with obstructive lung function. Respirology. 2015;20(4):587-93.

226. Landman GW, Kleefstra N, Groenier KH, Bakker SJ, Groeneveld GH, Bilo HJ, et al. Inflammation biomarkers and mortality prediction in patients with type 2 diabetes (ZODIAC-27). Atherosclerosis. 2016;250:46-51.

227. Li WJ, Chen XM, Nie XY, Zhang J, Cheng YJ, Lin XX, et al. Cardiac troponin and C-reactive protein for predicting all-cause and cardiovascular mortality in patients with chronic kidney disease: a meta-analysis. Clinics. 2015;70(4):301-11.

228. Viasus D, Del Rio-Pertuz G, Simonetti AF, Garcia-Vidal C, Acosta-Reyes J, Garavito A, et al. Biomarkers for predicting short-term mortality in community-acquired pneumonia: A systematic review and meta-analysis. The Journal of infection. 2016;72(3):273-82.

229. Jing X, Huang C, Zhou H, Li C, Fan L, Chen J, et al. Association between serum C-reactive protein value and prognosis of patients with non-small cell lung cancer: a meta-analysis. International journal of clinical and experimental medicine. 2015;8(7):10633-9.

230. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. J Intensive care medicine. 2020;46(5):846-8.

231. National Institute for Health and Care Excellence. Multimorbidity: clinical assessment and management (NG56) 2016 [Available from: https://www.nice.org.uk/guidance/ng56/evidence/full-guideline-pdf-2615543103].

232. Deelen J, Kettunen J, Fischer K, van der Spek A, Trompet S, Kastenmüller G, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nature Communications. 2019;10(1):1-8.

233. O'Kane MJ, Lopez B. Explaining laboratory test results to patients: what the clinician needs to know. BMJ. 2015;351:h5552.

234. Kravitz RL, Callahan EJ. Patients' perceptions of omitted examinations and tests - A qualitative analysis. Journal of general internal medicine. 2000;15(1):38-45.

235. Michiels-Corsten M, Donner-Banzhoff N. Beyond accuracy: hidden motives in diagnostic testing. Family Practice. 2017:cmx089-cmx.

236. Duddy C, Wong G. Explaining variations in test ordering in primary care: protocol for a realist review. 2018;8(9):e023117.

237. Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. Netherlands Journal of Medicine. 2007;65(5):167-77.

238. Merkebu J, Battistone M, McMains K, McOwen K, Witkop C, Konopasky A, et al. Situativity: a family of social cognitive theories for understanding clinical reasoning and diagnostic error. Diagnosis. 2020;7(3):169-76.

239. Lehman R. Sharing as the Future of Medicine. JAMA internal medicine. 2017;177(9):1237-8.

240. Pfeffer N, Laws S. 'It's only a blood test': what people know and think about venepuncture and blood. Social science & medicine (1982). 2006;62(12):3011-23.

241. Kljakovic M. Patients and tests - a study into patient understanding of blood tests ordered by their doctor. Aust Fam Physician. 2012;41(4):241-3.

242. Pomerantz A, Rintel ES. Practices for reporting and responding to test results during medical consultations: enacting the roles of paternalism and independent expertise. Discourse Studies. 2004;6(1):9-26.

243. Conrad P, Barker KK. The Social Construction of Illness: Key Insights and Policy Implications. Journal of Health and Social Behavior. 2010;51(1_suppl):S67-S79.

244. Mol A. What diagnostic devices do: the case of blood sugar measurement. Theoretical medicine and bioethics. 2000;21(1):9-22.

245. Rhodes LA, McPhillips-Tangum CA, Markham C, Klenk R. The power of the visible: the meaning of diagnostic tests in chronic back pain. Social Science & Medicine. 1999;48(9):1189-203.

246. van Ravesteijn H, van Dijk I, Darmon D, van de Laar F, Lucassen P, Hartman TO, et al. The reassuring value of diagnostic tests: a systematic review. Patient Educ Couns. 2012;86(1):3-8.

247. Petrie KJ, Muller JT, Schirmbeck F, Donkin L, Broadbent E, Ellis CJ, et al. Effect of providing information about normal test results on patients' reassurance: randomised controlled trial. BMJ. 2007;334(7589):352.

248. Hollnagel E. The ETTO Principle: Efficiency Thoroughness Trade-Off. . Surrey, England, Publishing, Ltd. 2009.

249. Veroff D, Marr, A., Wennberg, DE. Enhanced Support For Shared Decision Making Reduced Costs Of Care For Patients With Preference-Sensitive Conditions. 2013;32(2):285-93.

250. Epstein RM, Franks P, Shields CG, Meldrum SC, Miller KN, Campbell TL, et al. Patient-centered communication and diagnostic testing. Ann Fam Med. 2005;3(5):415-21.

251. Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns. 2014;94(3):291-309.

252. Joseph-Williams N, Edwards A, Elwyn G. Power imbalance prevents shared decision making. 2014;348:g3178.

253. O'Flynn N, Britten N. Does the achievement of medical identity limit the ability of primary care practitioners to be patient-centred?: A qualitative study. Patient Educ Couns. 2006;60(1):49-56.

254. Spiegelhalter D. Risk and uncertainty communication. Annual Review of Statistics. 2017;4:31-60.

255. Simpkin AL, Schwartzstein RM. Tolerating Uncertainty — The Next Medical Revolution? New England Journal of Medicine. 2016;375(18):1713-5.

256. van der Bles AM, van der Linden S, Freeman ALJ, Spiegelhalter DJ. The effects of communicating uncertainty on public trust in facts and numbers. Proceedings of the National Academy of Sciences. 2020;117(14):7672-83.

257. Sladakovic J, Jansen J, Hersch J, Turner R, McCaffery K. The differential effects of presenting uncertainty around benefits and harms on treatment decision making. Patient Educ Couns.99(6):974-80.

258. Meyer AND, Giardina TD, Bhise V, Singh H, Street RL, Menon S, et al. Patient perspectives on how physicians communicate diagnostic uncertainty: An experimental vignette study†. International Journal for Quality in Health Care. 2018;30(1):2-8.

259. Recchia G, Schneider CR, Freeman ALJ. How do the public interpret COVID-19 swab test results? Comparing the impact of official information about results and reliability used in the UK, US and New Zealand: a randomised, controlled trial. 2020:2020.12.04.20243840.

260. Boohaker EA, Ward RE, Uman JE, McCarthy BD. Patient notification and follow-up of abnormal test results - A physician survey. Arch Intern Med. 1996;156(3):327-31.

261. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson M. The global burden of diagnostic errors in primary care. BMJ Quality & Safety. 2017;26(6):484-94.

262. Litchfield IJ, Bentham LM, Lilford RJ, McManus RJ, Greenfield SM. Patient perspectives on test result communication in primary care: a qualitative study. Br J Gen Pract. 2015;65(632):e133-40.

263. Heshka JT, Palleschi C, Howley H, Wilson B, Wells PS. A systematic review of perceived risks, psychological and behavioral impacts of genetic testing. Genet Med. 2008;10:19-32.

264. Yu L, Li P, Yang S, Guo P, Zhang X, Liu N, et al. Web-based decision aids to support breast cancer screening decisions: systematic review and meta-analysis. Journal of comparative effectiveness research. 2020;9(14):985-1002.

265. Evans R, Edwards AGK, Elwyn G, Watson E, Grol R, Brett J, et al. 'It's a maybe test': mens' experiences of prostate specific antigen testing in primary care. Br J Gen Pract. 2007;57(537):303.

266. Beck S, Borutta B, Walter U, Dreier M. Systematic evaluation of written health information on PSA based screening in Germany. PloS one. 2019;14(8):e0220745.

267. Brush JE, Jr, Brophy JM. Sharing the Process of Diagnostic Decision Making. JAMA internal medicine. 2017;177(9):1245-6.

268. Whiting PF, Davenport C, Jameson C, Burke M, Sterne JAC, Hyde C, et al. How well do health professionals interpret diagnostic information? BMJ open. 2015;5(7).

269. Whiting PF, Rutjes AWS, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. Journal of clinical epidemiology. 2013;66(10):1093-104.

270. von Roon AC, Karamountzos L, Purkayastha S, Reese GE, Darzi AW, Teare JP, et al. Diagnostic Precision of Fecal Calprotectin for Inflammatory Bowel Disease and Colorectal Malignancy. Am J Gastroenterol. 2007;102:803.

271. Banks J, Hollinghurst S, Bigwood L, Peters TJ, Walter FM, Hamilton W. Preferences for cancer investigation: a vignette-based study of primary-care attendees. The Lancet Oncology. 2014;15(2):232-40.

272. Hamilton W, Round J. Identifying frailty in primary care. BMJ. 2017;358:j4478.

273. NHS England. Identifying frailty [Available from: https://www.england.nhs.uk/ourwork/clinical-policy/older-people/frailty/frailty-risk-identification/]

274. Bossuyt PMM, Bossuyt PMM, McCaffery K. Additional Patient Outcomes and Pathways in Evaluations of Testing. Medical Decision Making. 2009;29(5):E30-E8.

275. Rycroft-Malone J, Burton CR, Bucknall T, Graham ID, Hutchinson AM, Stacey D. Collaboration and Co-Production of Knowledge in Healthcare: Opportunities and Challenges. Int J Health Policy Manag. 2016;5(4):221-3.

276. O'Donoghue D. Ask three questions: The Health Foundation; 2011 [Available from: https://www.health.org.uk/blogs/ask-three-questions].

277. Centre for Academic Primary Care. The Consultation Open and Close (COAC) Study: University of Bristol; 2020 [Available from: http://www.bristol.ac.uk/primaryhealthcare/researchthemes/coac-study/].

278. Watson J, Salisbury C, Whiting P, Banks J, Pyne Y, Hamilton W. Added value and cascade effects of inflammatory marker tests in UK primary care: a cohort study from the Clinical Practice Research Datalink. BJGP. 2019;69(684):e470-e8.

279. Watson J, Jones H, Salisbury C, Whiting P, Banks J, Hamilton W. Use of multiple inflammatory marker tests in primary care: using Clinical Practice Research Datalink to evaluate accuracy. BJGP. 2019;69(684):e470-e8.

280. NIHR Applied Research Collaboration West. How useful are inflammatory marker tests in primary care? 2020 [Available from: https://arc-w.nihr.ac.uk/research/projects/how-useful-are-inflammatory-marker-tests-in-primary-care/].

281. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. BMJ. 2020;369:m1808.

282. Watson J, Whiting P. Coronavirus: how accurate are coronavirus tests? : The Conversation; 2020 [Available from: https://theconversation.com/coronavirus-how-accurate-are-coronavirus-tests-135972].

283. Watson J, McCormack J. Testing times with James McCormack and Jess Watson [Internet]. https://podcasts.apple.com/gb/podcast/testing-times-with-james-mccormack-and-jess-watson/id283916558?i=1000476788266: BMJ; 2020. Podcast

284. Kneale T, Watson J, Noble B. Flu and COVID-19 [Internet]. https://podcasts.apple.com/gb/podcast/flu-and-covid-19/id1474942018?i=1000490791453; 2020. Podcast

285. Watson J. False negatives, testing capacity and pheasants [Internet]. https://www.bbc.co.uk/programmes/m000jmm3: BBC Radio 4; 2020. Podcast

286. Watson J, Richter A, Deeks J. Testing for SARS-CoV-2 antibodies. BMJ. 2020;370:m3325.

287. Mytton O, McCarthy, N., Watson, J., Whiting, P. Interpreting a Lateral Flow SARS-CoV-2 Antigen Test. BMJ. 2021:[submitted].